HOUSEHOLD ARREARS IN UKRAINE: MICROECONOMETRIC EVIDENCE

by

Andriy Norets

A thesis submitted in partial
fulfillment of the requirements for the
degree of

Master of Arts in Economics

Kyiv-Mohyla Academy University

2002

Approved by _____
Chairperson of Supervisory Committee

_____

_____

_____

Program Authorized
to Offer Degree _____

Date _____

Kyiv-Mohyla Academy University

Abstract

HOUSEHOLD ARREARS IN UKRAINE: MICROECONOMETRIC EVIDENCE

by Andriy Norets

Chairperson of the Supervisory Committee: Professor Serhiy Korablin,
Institute for Economic Forecasting at the National Academy of Sciences of Ukraine

Although continual growth of household arrears for housing and utilities causes great concern among policymakers, there has been no empirical study on the micro level of this particular type of arrears in Ukraine and other transition countries. Applying logit, ordered probit, and Heckman's sample selection models, I conduct a microeconometric investigation of the determinants of household arrears in Ukraine. This empirical study is based on the data from a nationally representative household survey conducted by the State Committee of Statistics of Ukraine in 2000. It is shown that although the effect of the government program for the mutual cancellation of budgetary wage arrears and household arrears shows up in the data arrears to households are still important determinants of the presence and the size of household arrears. Estimation results indicate that while the poor are more likely to be in arrears for housing and utilities, the size of these arrears is positively related to income for those households that have nonzero arrears for housing and utilities. The explanation seems to be that the government program of housing subsidies creates strong incentives for poor people to comply, while the incentives for compliance for richer households are absent. Households headed by people with low level of education and, perhaps, little human capital are more likely to be persistent non-payers. Other things equal, old households seem to have higher payment discipline. Worklessness status of a household as a whole is a major determinant of household arrears. Policy implications concerning the issue are suggested in the paper.

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

ACKNOWLEDGEMENTS

I thank my thesis advisor, Dr. Hartmut Lehmann, for his guidance and invaluable suggestions regarding this research. I am greatly indebted to Dr. Wayne Thirsk, who kindly provided the data for the study, read all the drafts of this paper, and made insightful comments on my work. I express thanks to Dr. Ghaffar Mughul for commenting the latest draft of this paper. I am also grateful to Dr. Stefan Lutz, whose questions and suggestions on the presentations and drafts of the paper have helped me to improve this work. The author expresses his gratitude to Fiscal Analysis Office for the support of this research and the opportunity to acquaint policymakers with the results and policy implications of the thesis[1].

[1] The publication of a condensed version of this thesis is forthcoming in "Budget and Fiscal Review" of Fiscal Analysis Office of the Supreme Council Of Ukraine, June, 2002.

GLOSSARY

Arrear. It is a liability or a commitment that is due but not paid.

Arrears to household. Wage, pension, or social payment arrears to household.

Household arrears. Arrears of households for housing and utilities, if other is not specified.

Oblast. Subnational administrative unit in Ukraine. There are 24 oblasts in Ukraine.

Chapter 1

INTRODUCTION

Nonpayments have become a cultural phenomenon and a persistent disease in the economies of Newly Independent States of the Former Soviet Union. In Ukraine, household arrears for utilities and housing reached 3.5% of GDP in 2000 and continued to rise; the average period of indebtedness was equal to 12 months in 2000 (PADCO, 2001b). These facts cause great concern among policymakers.

The huge amount and prevalence of household arrears are a burden for the State Budget and a threat to sound operation of public utilities. The housing and communal sector has heavily worn out capital, the services supplied are of poor quality, and the sector is generally backward (developed countries as a share of GDP spend at least twice as much on housing as Ukraine does (UNDP, 1996)). Household arrears impede the much needed reform of the sector. Moreover, being a component of an interrelated arrears system typical for CIS countries, household arrears may aggravate the problem of arrears in general, thus hindering economic growth in the country (Morozov, Pinto, and Drebentsov, 2000). Even though household arrears are only one of many issues important for the housing and communal sector and Ukrainian economy in general, efficient legislative settlement of this problem is crucial for the development of the culture of contract enforcement which is an essential feature of a market-based economy.

Although considerable research has been devoted to arrears in general (Morozov, Pinto, and Drebentsov, 2000), wage arrears (Lehmann, Wadsworth, and Acquisiti, 1999), (Earl and Sabirianova, 2000), and (Gryshyna, 2000), and inter-enterprise arrears (Alfandari and Schaffer, 1996) and (Daianu, 1997), rather less attention has been paid to the problem of

household arrears in the economic literature. Previous attempts to clarify this issue in Ukraine ((PADCO, 1998), (PADCO, 2001), and (PADCO, 2001b)) are mainly descriptive. They are based on macro data and thus they cannot add much to an understanding of the phenomenon. Some works, such as (Bagratian and Gurgen, 1997), only allege the most probable causes of household arrears but do not support their conclusions with quantitative evidence. As a unique micro data set containing information on household arrears has recently become available, there is a niche for empirical research that concentrates on the determinants of household arrears at the micro level. Application of econometric tools for the analysis of arrears on the microlevel has already produced insightful results in studies of wage arrears, such as (Lehmann, Wadsworth, and Acquisiti, 1999). The microeconometric investigation of the determinants of household arrears should also be of great interest for policy makers who have to cope with this problem in Ukraine.

Wage, pension, and social payment arrears to households, poverty, and tariff reform as a result of which households started to pay almost full services costs instead of symbolic payment in Soviet times are recognized in the literature and public debates as the main causes of household arrears for housing and utilities. However, the actual situation with household arrears in Ukraine seems to be much more complex.

The government program of targeted subsidies for housing and utilities supporting the poor and the "implicit" government subsidy in the form of unenforceable contracts in the field and the absence of fines for late payments make the role of income and poverty in explaining household arrears ambiguous. This empirical study sheds light on the true relationship between household income and arrears for housing and utilities and provides indirect evidence about the effect of the housing subsidies program and the "implicit" subsidy on population payments compliance.

Wage and other arrears to households might not currently be an underlying cause of household arrears as the government introduced the program of the mutual cancellation of budgetary arrears to households and arrears for housing and utilities in 1999. Moreover, wage arrears and poverty have been falling since 1999, while household arrears have continued rising. The impact of arrears to households on household arrears is explored in the paper. Some evidence about the effectiveness of the program of the mutual cancellation of budgetary arrears to households and arrears for housing and utilities is also presented.

The effects of such interesting characteristics as worklessness status of the household as a whole, and age, sex, and education level of the household head on the presence, persistence, and the size of household arrears are also considered in the paper.

Quantitative evidence and its interpretation presented in the thesis serve to deepen the level of debates about the issue of household arrears, provide some indirect evaluation of the related government programs, and reveal the negative consequences of the absence of a well functioning economic law punishing the breach of contract.

The organization of this paper is as follows. Chapter 2 describes the history of household arrears in Ukraine and the main legislative regulations governing the behavior of customers and providers of utilities in regard to household arrears. In addition, this chapter investigates the determinants of household arrears at the regional level using graphical analysis and the results of the panel data model estimation. Chapter 3 contains a literature review section discussing the general framework for the analysis of arrears and the possible role of household arrears in this framework. It also lists possible causes of household arrears proposed in previous research. Several studies on wage arrears used as prototypes for building empirical models in this paper are briefly summarized in the section. Chapter 4 presents the data description and

hypotheses on the role of income, wage, pension, social payment arrears, worklessness status of the household as a whole, and other characteristics in explaining household arrears. Chapter 5 includes econometric specifications of the logit models used for the analysis of the determinants of the presence and persistence of household arrears and ordered probit and tobit II models used for the analysis of the determinants of the size of household arrears. A discussion of the estimation results is also given in this chapter. Chapter 6 concludes with a summary of the results and policy implications. The descriptions of the datasets, the listing and log file of the Stata computer program, and the summary of the estimation results for microeconometric models are given in appendices.

Chapter 2

HOUSEHOLD ARREARS IN UKRAINE

2.1 The history and regulations of household arrears in Ukraine

Until 1995, the housing and utilities sector was mainly financed from the State Budget of Ukraine. Customers were to pay only 4% of services costs in 1994; the other part of the costs was covered by the government subvention which comprised 8% of total budget expenditures (PADCO, 2001b, p. 1). The purpose of the reform of tariffs for housing and utilities initiated in 1995 was to improve the services quality which was very low, decrease the budget burden which increased due to rise in energy prices, and to create incentives for efficient use of resources by suppliers and customers. The reform of tariffs for housing and utilities was an important step in moving towards a market-based economy. As a result of this reform, the portion of customers' payment for housing and communal services was dramatically increased to 60% of provider costs in 1996. The increase in prices of housing and utilities, falling real income of the population, and the growth of wage arrears have all contributed to the rise in household arrears for the services consumed.

The program of subsidies for housing and utilities was introduced in 1995 in order to assist those poor families unable to pay. Thus, Ukraine refused to subsidize housing and utilities for the whole population and provided targeted aid only for the needy. According to the Cabinet of Ministers' Resolution "On provision of subsidy for compensation of expenditure on payment for housing and communal services," from the 1st of May 1995 households were to pay for housing and communal services not more than 15% of their monthly income. The remaining part of the payment for the services automatically became a liability of the local budget. Later, quotas on consumption of electricity and gas on which the subsidy can be granted were

defined in the legislation. In 1998, the share of household average income which is to be paid for housing and utilities was increased to 20% (some categories of households were still to pay 15 %). Nearly one fifth of the Ukrainian population got the subsidy during 1997-2000 (PADCO, 2001b, p. 31). In PADCO's expert opinion, the program of housing subsidies was an important factor in improving household compliance in recent years.

According to the Ukrainian Ministry of Labor and Social Policy, the tariff reform and the program of housing subsidies decreased the budget burden by approximately UAH 2 billion in 1999. This is recognized as a great achievement of the reform, which succeeded in assisting the poor and decreasing the budget burden despite numerous depressing forecasts.

In 1995, the Cabinet of Ministers of Ukraine suspended fines on delays in payment for housing and communal services. This measure could be defended at the time when real population income was dropping and prices for housing and utilities were flying up by 1000-3000 % during the year. However, now the absence of fines aggravates the problem of household arrears and does not seems to be well justified.

The resolution of the Supreme Council of Ukraine "On payment for housing and communal services by population of Ukraine" declared in 1999 that the Cabinet of Ministers of Ukraine, local authorities, enterprises and organizations irrespective of ownership type were prohibited to cut off living quarters from energy, heat, water, and gas and evict people from their dwellings for the reason of nonpayment if these people had wage or pension arrears and did not get the subsidy. By adopting this resolution, the Supreme Council of Ukraine exceeded its authorities as determined by the Ukrainian Constitution. Therefore, some public utility providers do not consider it and sometimes cut off nonpayers from energy, heat, water, and gas. Nevertheless, the resolution is likely to have crucially contributed to the initial deterioration of household payments discipline.

After the period of striking changes, the situation stabilized. Since 1997, the rate of compliance of population has remained almost constant at near 77%. By the end of 2000, cumulative household arrears amounted to UAH 6.2 billions and the average period of indebtedness reached 12 months.

As the situation with wage and social payment arrears has recently improved and the targeted aid program of housing subsidies has succeeded in assisting the poor to pay for the services consumed, more effort is being made now to fight against persistent non-payers. There was an attempt to tackle the problem in 2000. The Supreme Council of Ukraine considered adopting the Law "On collecting arrears for housing and communal services". This Law offered to create special local commissions to deal with each case of household arrears in deciding whether an indebted household had a good excuse such as wage arrears for having such debt. If the commission does not find a good excuse, the arrears, along with a fine, must be collected from the household. The Supreme Council of Ukraine postponed the consideration of the Law justifying its decision by the existence of wage, pension, and social payment arrears. It seems, however, that there is no logical link between the decision and the content of the proposed Law. Moreover, budgetary arrears to households cannot currently be a cause of household arrears for housing and utilities since in 1999 the Cabinet of Misters of Ukraine designed a procedure for the mutual cancellation of household arrears and pension, social payment, and wage arrears from the budget.

The last resolution concerning the problem developed by the Supreme Council of Ukraine in 2001 forgives arrears for housing and communal services for those households that have income lower than some predefined level. This pre-election proposal is purely populist one and its adoption would further discourage payment discipline. As the resolution was not signed by the President of Ukraine and there have been new elections to Ukrainian Parliament, the chances to find an efficient solution to the problem of household arrears in Ukraine have improved.

2.2 Evidence from aggregate data

Although aggregate data can uncover little about the determinants of household arrears at the micro level, the analysis of these data may shed some light on the dynamics of the problem, which cannot be investigated on the micro level because only a cross-sectional household level dataset is available. Moreover, the results of the analysis can be useful for formulating some preliminary hypotheses for the study on the micro level.

Figure 2.1 plots the aggregate stocks of household and wage arrears in Ukraine for the period of 1998:12-2001:11. Since August 1999 the stock of wage arrears has been falling, while the stock of household arrears has continued rising. This picture is similar for the majority of Ukrainian oblasts.

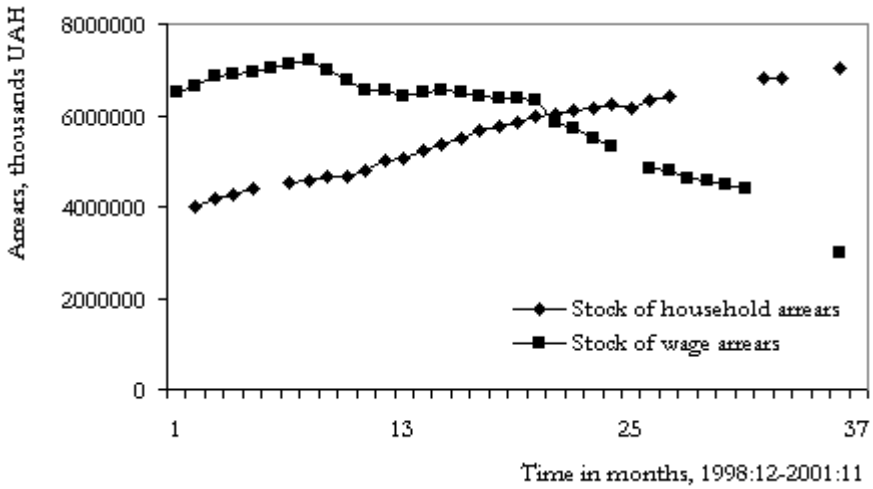

Figure 2.1. The stocks of wage arrears and household arrears (source: express-reports of Ukrainian State Committee of Statistics)

After August 1999 real income per capita has also been increasing in Ukraine. The fact that household arrears are growing while income is rising and wage arrears are decreasing should cause concern among those who explain household arrears by wage arrears and poverty alone. In fact, the correlation between income and household arrears at the regional level is positive, rather than negative. The analysis of the oblast level data elucidates the problem in more detail.

Each point on the figure 2.2 represents a region in Ukraine (regions include 24 Ukrainian oblasts, Crimean Autonomy, and the cities of Kyiv and Sevastopol). The region's average stock of household arrears per capita in 2000 is measured on the vertical axis. The region's income per capita in 2000 is measured on the horizontal axis. Visual inspection of figure 2.2 indicates that there is a positive correlation between the size of household arrears and income per capita on the regional level. Richer oblasts have higher stock of household arrears per capita. The outlier corresponds to Kyiv, which is the capital of Ukraine. Income per capita is much higher in Kyiv than in the other regions as the concentration of economic activity is very high there. The results are the same for 1999.

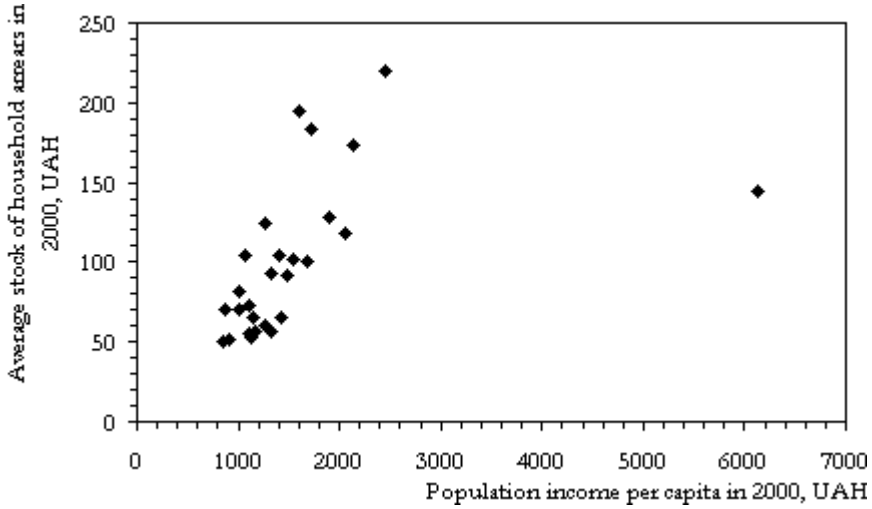

Figure 2.2. Average stock of household arrears per capita and yearly population income per capita by oblasts in 2000 (source: express-reports of State Committee of Statistics of Ukraine and 2000 statistical yearbook of Ukraine)

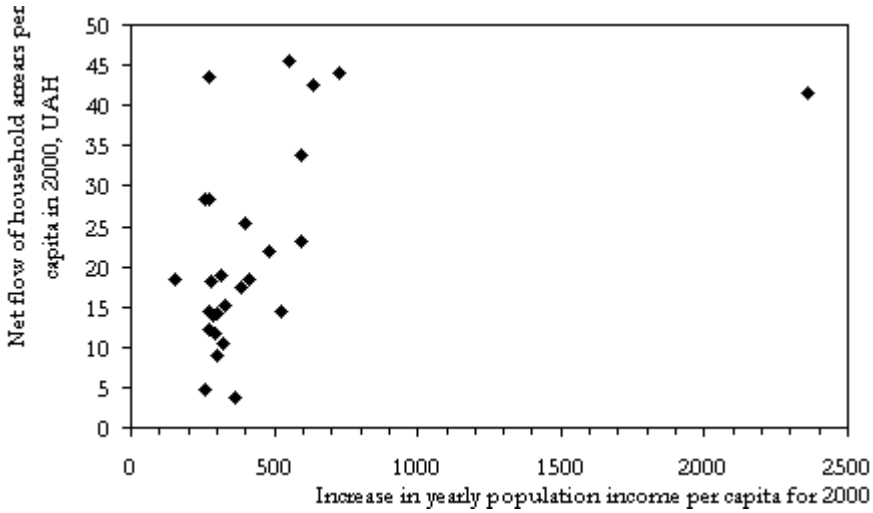

Figure 2.3. Net flow of household arrears and change in yearly population income per capita by oblasts in 2000 (source: express-reports of State Committee of Statistics of Ukraine and 2000 statistical yearbook of Ukraine)

Figure 2.3 presents the graph of the net yearly flow of region's household arrears per capita against the increase in region's yearly income per capita in 2000. It can be seen from the figure 2.3 that the larger is an increase in income per capita in an oblast, the larger is an increase in household arrears per capita in this oblast. The same picture is observed in 1999. Analogous investigation of the relationship between household and wage arrears is presented below.

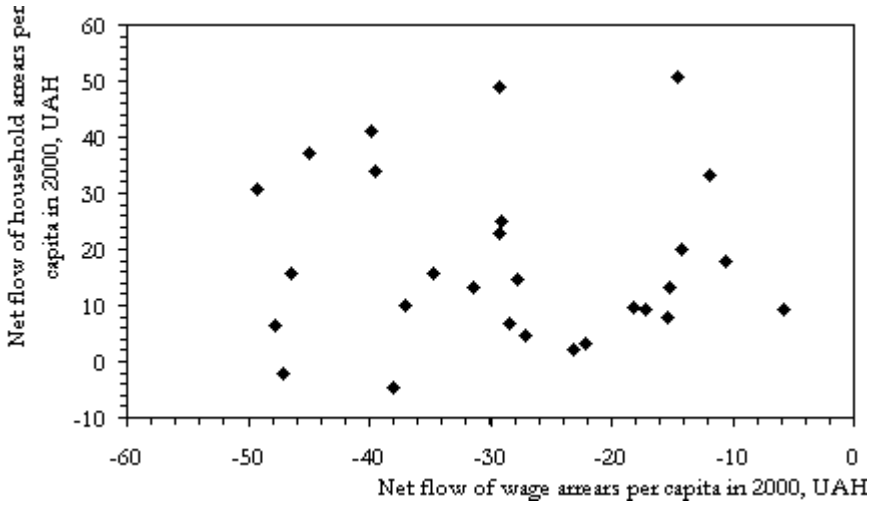

Figure 2.4. Net flows of household and wage arrears per capita in 2000 by oblasts (source: express-reports of State Committee of Statistics of Ukraine)

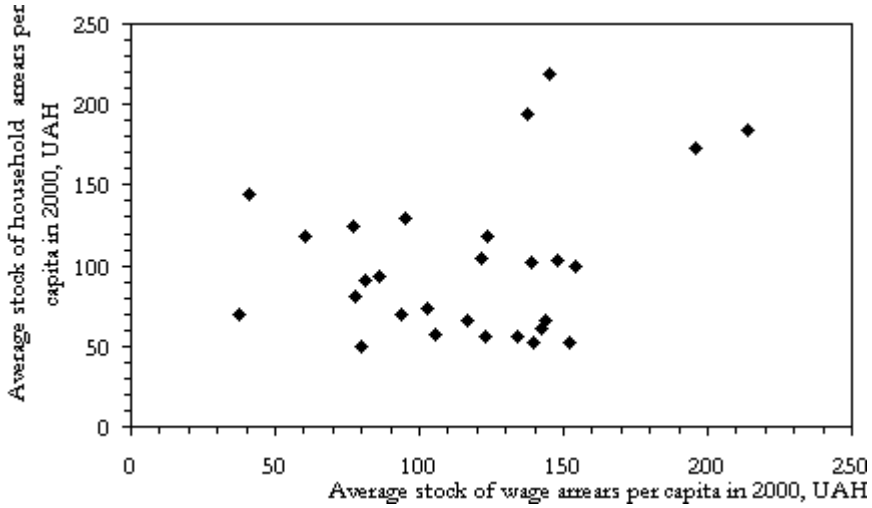

Figure 2.5. Average stocks of wage arrears and household arrears per capita in 2000 by oblasts (source: express-reports of State Committee of Statistics of Ukraine)

Figures 2.4 suggests that there is a weak link between flows of household arrears and wage arrears on the oblast level. However, as figure 2.5 shows there are some signs of positive correlation between the stocks of household and wage arrears per capita, which mainly stem from the four upper points on the graph (Dnepropetrovska, Donetska, Kharkivska, and Luganska oblasts).

In order to further investigate the relationship between household arrears, income, and wage arrears on the aggregate level we use regression analysis. A fixed effects panel data model should be used since the sample includes the whole population and regions seem to be heterogeneous in Ukraine. It is assumed that the net flow of household arrears per capita in the region *i* at year *t*, Δha_t^i, depends linearly on the net flow of wage arrears per capita, Δwa_t^i, the increase in income per capita, $\Delta income_t^i$, and the share of urban population in total population, $urbanshare_t^i$, in the region *i* at year *t*:

$$\Delta ha_t^i = \alpha * \Delta wa_t^i + \beta * \Delta income_t^i + \gamma * urbanshare_t^i + \varepsilon_t^i \qquad (2.1)$$

where ε_t^i is an error term, $\alpha, \beta, and \gamma$ are coefficients to estimate. The choice of the independent variables and the equation specification is mainly determined by the data available. Fixed effects that control for possible heterogeneity of the regions are differenced out in the presented specification.

Available panel data include 54 individual observations: observations on 27 Ukrainian regions (24 oblasts, Crimean autonomy, and the cities of Kyiv and Sevastopol) in 1999 and 2000. The estimation was performed by GLS; cross section weights were estimated in preliminary regression with equal weights and then applied in weighted least squares in the second round. The results of estimation are presented in table 2.2.

Table 2.1. Estimation results for panel data model

Dependent variable: net flow of household arrears per capita		
Independent variables:	Coefficient	p-value
Net flow of wage arrears per capita	0.1182	0.0057
Change in yearly population income per capita	0.0106	0.0000
Share of urban population in total population	27.3984	0.0000
R-squared = 0.45, Prob(F-statistic) = 0.0000		

All the coefficients are significant and the F-statistic is highly significant. Adding the dummy variable for Kyiv (Kyiv was the outlier, figures 2.2 and 2.3) does not qualitatively change the estimation results. The model is not tested for heteroscedasticity as it is estimated by WLS. There is little sense in testing for autocorrelation as there are only two points in time. There are no available candidates for omitted variables. Thus, the model was not subjected to tests for omitted variables. Whatever the case, the results of this estimation are preliminary in nature, as this is a macrolevel analysis which may poorly reflect the situation on the household level.

Modeling results show that there is a positive relationship between household arrears and income on the regional level even after we control for some other factors. Controlling for the effects of income and the share of urban population provides evidence that the relationship between household arrears and wage arrears is positive, which was not obvious from the visual inspection.

The household level study corroborates the findings of this section and also considers a few other determinants of household arrears. Possible explanations and policy implications of the findings are given in Chapter 5 with the discussion of the results of microeconometric investigation.

Chapter 3

HOUSEHOLD ARREARS IN ECONOMIC LITERATURE

Different types of arrears in a CIS country seem to be highly interrelated. Therefore, in order to investigate certain type of arrears one should understand the whole system in general. A universal framework for the analysis of arrears in the Russian Federation is presented in Morozov, Pinto, and Drebentsov, (2000). Authors stated that "nonpayments intensified and spread as a result of inconsistency between macroeconomic and microeconomic policy". According to them, the goal of macroeconomic policy was rapid stabilization accomplished by tightening credit and fixing exchange rate, while the implicit microeconomic goal was "to maintain social safety net by continuing to subsidize enterprises thereby encourage them to remain in operation". Enterprises received these implicit subsidies by not paying taxes and energy bills. In turn, energy sector monopolies, Gasprom and RAO UES, compensated these losses from unpaid services by running tax arrears and participating in mutual settlements.

The situation described by the authors for Russia is similar to the one in Ukraine. Although household arrears for housing and communal services were not considered by the authors, we can hypothesize about their role in the scheme. Household arrears may be an implicit subsidy from the government to households compensating for wage arrears and the rapid rise in prices for housing and utilities. This subsidy is mediated through the energy sector and communal services providers. How inefficient the distribution of this implicit subsidy is a question for our empirical investigation.

The study of inter-enterprise arrears by Daianu (1997) also presents an insightful view of arrears. The author sees the true cause of arrears in the real side of the economy. In a transition country, the magnitude of the arrears is

determined by the size of resource misallocation. "Arrears … influence the system in an ambivalent way: they seem to operate as a self-protecting device against the pressure for change … at the same time, they can slow down dangerously the speed of restructuring … by relaxing financial discipline". In the author's opinion, "setting up a clearing house, securitization of inter-enterprise credits are only temporary, or partial solutions; they do not focus on the main primary causes of arrears which … are the size of required structural adjustment". Thus, anti-arrears measures cannot be a one-shot policy. Daianu argues that as privatization and move towards market-based economy proceeds inter-enterprise arrears will decline and lose their deleterious effect on the economy. Analogous argument may apply to the problem of household arrears. However, it seems that the main stages of required adjustment in the housing and communal sector of Ukraine have already occurred. Customers are to pay for almost the whole services value. The size of government subventions for the sector has fallen considerably. Our study of the determinants of household arrears is intended to show that just a little political will is required to introduce some one-shot policies such as fines to finalize restructuring of housing and communal sector efficiently and rapidly and to solve the problem of household arrears in Ukraine.

A view of the causes of household arrears for housing and communal services is expressed by Bagratian and Gurgen (1997). The authors argue that "slippages in payment by households reflected, for the most part, a faster growth in the tariffs for gas, heating, and electricity than in household incomes … collection efforts were also weakened by difficulties in measuring energy consumption by households, given shortage of meters … this may also have heightened the unwillingness to cut off delinquent customers". The research also emphasizes political and social reasons for household arrears. Before elections in particular, the government discouraged cutting off gas, electricity, etc. for slippages in payment, since it had budgetary arrears on wages and pensions. This study presents no empirical confirmation for these

statements about main causes of arrears and thus, our tests may shed light on the issue using quantitative evidence.

The studies of wage arrears at the micro level, such as (Lehmann, Wadsworth, and Acquisiti, 1999), (Earl and Sabirianova, 2000), and (Gryshyna, 2000), are very instructive for our work since they demonstrate how microeconometric tools can be applied for the analysis of arrears. Research on wage arrears addresses two main questions: Why do firms engage in regular late payment of wages to their workers? Why do workers tolerate wages in arrears? Analogous questions can also be asked about household arrears for housing and utilities: Why do people not pay for the services consumed? Why have the government tolerated this practice for the last eight years? To answer the questions about wage arrears in Russia Lehmann, Wadsworth, and Acquisiti (1999) applied the whole range of econometric tools. Their study was based on large labor market surveys. They used cross tables to look at the incidence and persistence of wage arrears by region, sector, industry, etc. To reveal the factors which determine the probability of having wage in arrears the authors employed probit and random effects probit models. For the analysis of the determinants of the persistence of wage arrears ordered probit model was estimated. Estimation results of tobit model identified the factors that determine the sum of wages in arrears. Multinomious logit model was used for studying the mobility of workers affected by wage arrears. A number of interesting hypotheses were examined. Earl and Sabirianova (2000) and Gryshyna (2000) also applied analogous microeconometric models to check some additional hypotheses.

Since 1996, the consulting firm PADCO International has monitored data on arrears of payment for housing and communal services in Ukraine and has published series of annual analytic reports on this issue. Although the reports have mainly a descriptive character they are the only attempt at research specialized in the area and, thus, can be used for the formulation of preliminary hypotheses on causes of the household arrears. PADCO

(PADCO, 1998, p.4) suggests that the size of payment arrears for housing and communal services depends on at least four factors: 1) the size of wage, pension, and social payment arrears; 2) accounting precision and collection speed of payments (the more transparent and efficient accounting system the less the size of arrears); 3) household income level (poor rayons have comparatively large arrears); 4) quality of housing and communal services (the poorer the services the less the propensity to pay for them). The reports provide no quantitative evidence on these possible causes of household arrears as well as no evidence on the magnitude of their influence.

While there is no data available to test the importance of the second and the fourth factors in explaining household arrears, the first and the third factors are considered in our research. Contrary to the PADCO report, the effect of level of household income seems to be ambiguous. As we already mentioned, the government housing subsidies for the needy and the "implicit" subsidy in the form of unenforceable contracts and the absence of fines for late payments distort the proposition that household arrears are a destiny of the poor. The analysis of the regional level data presented in the section 2.2 also does not support this proposition. Another argument against poverty as the main cause of arrears is that the more the income the more the opportunity costs of time spent on conducting payment and since there is no strict enforcement for on time payments, people with high 'shoe leather' costs, that is high earned income, may chose to pay more rarely than others. In PADCO expert's opinion (PADCO, 1998, p.4), the absence of penalties and fines for late payment and nonpayment significantly aggravate the problem of arrears. Empirical investigation can help to check such judgements.

The coincidence of wage arrears and household arrears for housing and communal services is also not obvious as the program of the mutual cancellation of budgetary arrears to households and arrears for housing and utilities was introduced in 1999. Another puzzle for a proponent of wage arrears as a cause of household arrears is that wage arrears and poverty have

been falling since 1999, while household arrears have continued rising. Thus, the role of arrears to household is a subject for empirical tests in our work.

Analysis of the data on household arrears, which are monitored by PADCO for city Kharkiv (PADCO, 1998, p.10; PADCO, 2001, p.21), shows that while the majority of population pays for housing and communal services with a constant delay, 43-44% of the population are persistent nonpaying customers, whose average debt reached UAH 1054 per family in 2000. Such analysis of observations by PADCO can help us interpret the indicators of arrears in the cross-section data set available and stimulate to investigate characteristics of persistent nonpaying customers.

This short literature review shows that economic research supported by empirical work in the area of household arrears for housing and utilities is of great interest since there are some hypotheses about possible causes and characteristics of the phenomenon in the economic literature on arrears; however, there has been no attempt at deep empirical analysis of the issue.

Chapter 4

HYPOTHESES AND DATA

4.1 Data description

Since 1999, the State Committee of Statistics of Ukraine conducts a quarterly household survey, which covers more than 12000 of households. Starting from 2000, this nationally representative survey includes questions on household arrears for the services consumed as well as on arrears to households. The dataset from the 2000 rounds of the survey has recently become available. The dataset consists of two parts: household level data which include 9318 observations and individual level data which include 25133 observations. Household level dataset contains information specific to the household in general such as data on expenditure and household composition. Individual level dataset contains information about each household member for each of the households from the former dataset. Individual level data includes information on education, socio-economic status, income, etc. Detailed description of both datasets is given in Appendix A.

The data on household arrears for housing and communal services and arrears to household are available only in categorical form. Six categories for the size of the household arrears and seven categories for the size of the arrears to household are shown in the tables 4.1 and 4.2.

Table 4.1. Representation of the data on household arrears.

Category	0	1	2	3	4	5
The size of arrears, UAH	0	Under 100	100,1-300	300,1-500	500,1-1000	Above 1000
Proxy for the size of arrears, UAH	0	50	200	400	750	2000, (1500, 3000 - for sensitivity analysis)

Table 4.2. Representation of the data on arrears to households.

Category	0	1	2	3	4	5	6
The size of arrears, UAH	0	Under 200	200,1-300	300,1-500	500,1-1000	1000,1-2000	Above 2000
Proxy for the size of arrears, UAH	0	100	250	400	750	1500	3000, (2500, 4000 - for sensitivity analysis)

The median of the interval to which the size of household arrears falls divided by the number of household members is used as a proxy for the size of household arrears per household member in our empirical study. For the last category the proxy is not defined, thus we use several alternatives in econometric model estimation to perform the sensitivity analysis. Analogously, the proxy for the share of arrears to household in total household income is constructed. Other variables that are used in econometric models are discussed in the next section.

4.2 Hypotheses

4.2.1 Role of poverty and income level

The role of poverty and income level in explaining household arrears is far from obvious. The reason for that is government regulations of the issue. It could be natural to expect poor families to be in arrears and rich families to pay regularly. However, the targeted housing subsidies program creates high incentives for the poor to comply because if people pay for the services consumed and gradually repay old debts they are eligible for the subsidy under which they are to pay for housing and utilities no more than 20% of their income. On the contrary, the rich are not motivated by the program at all. Moreover, high 'shoe leather' costs may worsen payment discipline of the rich. Households should pay for housing and utilities once per month. Payment procedure requires visiting a bank and standing in a queue. Therefore, it takes time and implies opportunity costs. The benefits of regular payments are the absence of annoying and threatening remindings and fines.

As fines for arrears were abolished in Ukraine, benefits from regular payment seem to be negligible and independent from household income. Thus, the higher an individual has income the more his or her opportunity costs for paying regularly which can be called 'shoe leather' costs. Empirical inquiry will shed light on the validity of the above arguments.

Cash income of the household reported in the survey does not completely determine the ability of the household to pay for the services. There are several reasons for that. The first one is that "private or 'inter-household' transfers may significantly change consumption possibilities" and such informal support networks play an important role in a transition society experiencing a rise of poverty (Coudouel, McAuley, and Micklewright, 1997). The other reason is that households support themselves by growing vegetables, fruits, and livestock, which also changes household consumption possibilities. In order to identify the role of poverty in explaining household arrears more precisely we consider both mentioned factors in our empirical investigation in addition to cash income received by the household. Thus, under the household income we understand total resources available to the household that are equal to the sum of cash and in-kind earned income, value of household production, and value of 'inter-household' transfers.

Looking at descriptive statistics calculated from the sample available may help to initially characterize the relationship between income and household arrears. The listing and log of the Stata computer program which calculates descriptive statistics for the cross-tables in this section are presented in Appendices B and C. For more detail, such as standard deviations, maximal and minimal values, see log file in Appendix C.

Table 4.3 gives the sample means of income per household member for two subsamples. The first subsample contains observations on the households that do not have arrears for housing and communal services; the second subsample contains all remaining observations.

Table 4.3 Sample mean of household income per member

Subsamples	Sample mean of income per member, UAH	Number of observations in the subsample
Household with arrears	1790	2840
Households without arrears	2068	6478
Total	1978	9318

On average, households with arrears have lower income per member than those without arrears. Figure 4.1 shows the sample means of income per household member calculated for the subsamples corresponding to categories of the size of household arrears per household member. On the diagram, these categories are represented by the proxies for the size of household arrears per household member (the details on the proxy calculation are discussed in the section 4.1).

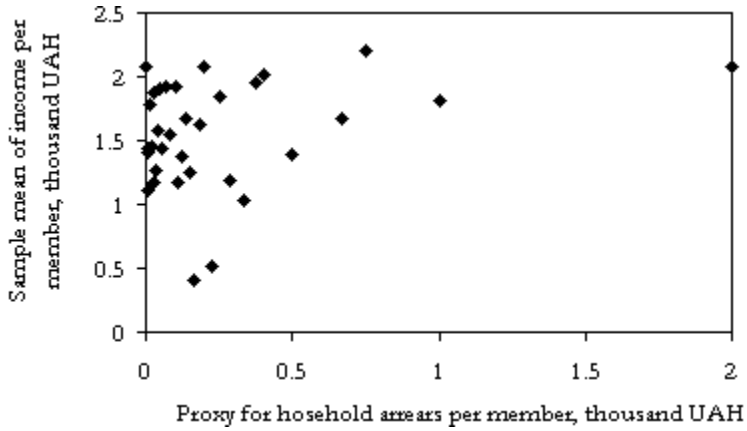

Figure 4.1. Sample means of income per household member

Figure 4.1 suggests that the household income per member is positively related to the size of household arrears (the correlation coefficient between proxy for the size of households arrears per household member and the sample mean of income per household member for corresponding subsample is equal to 0.3). The same conclusion is drawn from the analysis of the macro data in the section 2.2. The results of econometric modeling presented in

chapter 5 corroborate this finding. An explanation for this and corresponding policy implications are given in the next chapter.

4.2.2 Role of arrears to household

The presence of wage, social payment, and pension arrears is supposed to be an important determinant of household arrears (PADCO, 1998) and is suggested as an excuse for not reintroducing fines for nonpayments by Supreme Council of Ukraine. Arrears to household may contribute to the problem through two channels. First, they could lead to poverty. Second, they might provide a psychological excuse for nonpayment. Empirical tests should provide evidence on how important these two channels may be. Although the government program for the mutual cancellation of budgetary arrears to households and household arrears for utilities and housing may weaken the relationship between them, descriptive statistics given in table 4.4 suggest that the households with arrears for communal and housing services more often and to a greater extent suffer from wage arrears.

Table 4.4 Descriptive statistics for wage, social payment, or pension arrears by subsamples.

Subsamples	Proportion of households that have wage, social payment, or pension arrears	Mean of share of wage, social payment, or pension arrears in total household income	Number of observations in the subsample
Household with arrears	0.303	0.062	2840
Households without arrears	0.217	0.036	6478
Total	0.245	0.044	9318

4.2.3 Role of worklessness on household level

The worklessness status of a household as a whole may reflect the extent of social distress and exclusion for this household and, consequently, may determine the ability or willingness to pay for the services consumed.[2] The fact that unlucky labor market experience is concentrated in a household (not

[2] The idea of looking at the worklessness level of the household came from my thesis advisor Professor Hartmut Lehmann

one but several household members are unemployed) may play a very important role in explaining the presence and the size of arrears of this household. After Gregg and Wadsworth (2001), we classify households by their worklessness level in the following way: "a workless household is observed when all the adult occupants are out of work," "an all-work household is observed when all the adult occupants are in work," and a mix-work household is observed when some adult occupants work but some other do not. Three corresponding dummy variables are used in the econometric analysis of the next chapter. They were constructed from individual level dataset and added to the household level dataset. SQL computer program, which is used for creating the household worklessness level variable, is presented in the Appendix B. A pensioner household, a household in which all adults are of pension age, is considered as separate household type. Students are treated as if they are in work. Housekeepers are treated as if they are out of work. Sample proportions of household of four types are presented in the table 4.5.

Table 4.5 Descriptive statistics for household worklessness status.

Subsamples	Proportion of pensioner households	Proportion of all-work households	Proportion of mix-work households	Proportion of workless households	Number of observations in the subsample
Household with arrears	0.147	0.506	0.238	0.109	2840
Households without arrears	0.341	0.400	0.176	0.083	6478
Total	0.278	0.434	0.196	0.092	9318

As can be seen from the table, the proportion of pensioner households for the households without arrears is much higher than for the households with arrears. This may indicate that old age households have higher payment discipline than young ones. The effect of worklessness level is not obvious

from the table and it is considered in the econometric study in the next chapter.

4.2.4 Other household characteristics

A few household characteristics, such as age, gender, education level, or field of household head, are supposed to be the determinants of household arrears. Thus, old age households accustomed to strict discipline during soviet 50's and 60's may be more disciplined in paying for services than younger ones with similar other characteristics. Well-educated and progressive households may be less risk averse and less afraid of fine reintroduction and thus, may have looser payment discipline. On the other hand, people with high human capital are easily able to pay for the services consumed. Households headed by a woman might have high arrears due to low earned income. In the next chapter we present and discuss empirical evidence that supports or refutes the importance of a range of household characteristics in explaining household arrears for housing and utilities.

C h a p t e r 5

MICROECONOMETRIC STUDY

5.1. Models and methods

5.1.1. Binary choice model

A binary choice model can be used for the analysis of the incidence of household arrears. In a binary choice model the dependent variable is limited to two values. In our context, if the household *i* has arrears then $y_i=1$, otherwise $y_i=0$. Then it is assumed that the probability of having arrears for the household *i* is a function of a vector of individual characteristics x_i and a vector of parameters β :

$$P(y_i = 1 \mid x_i) = F(x_i'\beta) \tag{5.1}$$

where a probability distribution function is a natural choice for *F(.)* as $P(y_i = 1 \mid x_i)$ should lie inside *(0,1)*. If a logistic probability distribution function is used then the model is called logit; if a standard normal probability distribution function is used then the model is called probit or normit.

If linear regression $y_i = x_i'\beta + \varepsilon_i \; and \; E(\varepsilon_i \mid x_i) = 0$ is used instead of a binary choice model, the following problems will appear: error terms are highly non-normal and heteroscedastic; fitted values of y_i might not lie between 0 and 1 (see, for example, (Verbeek, 2000, section 7.1). A binary choice model can solve these problems and, in addition, allow for a non-linear relationship between $P(y_i = 1 \mid x_i)$ and $x_i'\beta$.

Maximum likelihood method is used for estimation of vector of parameters β in (5.1). The Huber-White-sandwich estimator of variance is used in place of the traditional calculation to avoid the problem of heteroscedasticity (for

details on obtaining robust variance estimates see Stata User Manual, [U]23.11). The results of estimation for two logit specifications and their discussion are presented in sections 5.2.1 and 5.2.2.

5.1.2. Ordered probit model

Ordered probit model is used for the analysis of the factors that determine the size of arrears that a household has. This model is chosen since the data on household arrears are only reported to lie in certain ranges. It is assumed that

$$
\begin{aligned}
& y_i^* = x_i'\beta + \varepsilon_i \\
& \textit{no intercept} \\
& y_i = j \ \textit{if} \ \mu_{j-1} < y_i^* \le \mu_j \\
& \mu_0 = -\infty, \ \ \mu_m = \infty \\
& \varepsilon_i \to N(0,1)
\end{aligned}
\tag{5.2}
$$

where y_i^* is the actual size of arrears that household i has, x_i is a vector of individual characteristics, β is a vector of parameters, y_i is observed value ($y_i=1,2,\ldots,m$), m is the number of ranges reported, μ_j's are bounds of ranges. The probability that y_i^* lies in a range j can be easily defined under the assumptions made about the distribution of the error term:

$$P(y_i = j \mid x_i) = \Phi(\mu_j - x_i'\beta) - \Phi(\mu_{j-1} - x_i'\beta) \tag{5.3}$$

where $\Phi(.)$ is standard normal probability distribution function.

Maximum likelihood method is used for estimation of vector of parameters β in (5.2). The Huber-White-sandwich estimator of variance is used in place of the traditional calculation to avoid the problem of heteroscedasticity (for details on obtaining robust variance estimates see Stata User Manual, [U]23.11). The results of estimation for probit specification and their discussion are presented in section 5.2.2.

5.1.3. Heckman's sample selection model

The sample available predominantly consists of the observations on the households that do not have arrears (near two thirds of the sample). Thus, categories other than "zero household arrears" may be underrepresented in the ordered probit analysis. Therefore, Heckman's sample selection model is used alternatively to ordered probit model for the analysis of the factors that determine the size of household arrears. The model consists of two equations. In the context of this paper (exposition here is based on Verbeek, 2000, section 7.4.1), the first equation, the arrears size equation, is

$$y_{1i}^* = x_{1i}'\beta_1 + \varepsilon_{1i} \tag{5.4}$$

where y_{1i}^* is the size of arrears per household member that household i has, x_i is a vector of individual characteristics that determine the size of household arrears, β_1 is a vector of parameters. y_{1i}^* is not observed for those households that do not have arrears. The second, a binary choice type equation is specified to describe whether a household has arrears:

$$y_{2i}^* = x_{2i}'\beta_2 + \varepsilon_{2i} \tag{5.5}$$

where y_{2i}^* is an unobserved variable that can be interpreted as a propensity not to pay, x_i is a vector of individual characteristics that determine whether a household has arrears, β_2 is a vector of parameters. The following observation rule is used:

$$\begin{aligned} &y_{1i} = y_{1i}^*, \;\; y_{2i} = 1 \; if \; y_{2i}^* > 0 \\ &y_{1i} \; is \; not \; observed, \;\; y_{2i} = 0 \; if \; y_{2i}^* \le 0 \end{aligned} \tag{5.6}$$

The model is completed by distributional assumptions about error terms:

$$\begin{pmatrix} \varepsilon_{1i} \\ \varepsilon_{2i} \end{pmatrix} \rightarrow N\left[\begin{pmatrix} 0 \\ 0 \end{pmatrix}, \begin{pmatrix} \sigma_1^{\,2} & \sigma_{12} \\ \sigma_{12} & 1 \end{pmatrix} \right] \tag{5.7}$$

Under the assumptions (5.7), it can be shown (see Greene, 2000, section 20.4) that:

$$E(y_{1i} \mid y_{2i} = 1) = x_{1i}'\beta_1 + \sigma_{12}\frac{\phi(x_{2i}'\beta_2)}{\Phi(x_{2i}'\beta_2)} \tag{5.8}$$

where $\Phi(.)$ is standard normal probability distribution function, $\phi(.)$ is standard normal probability density function, $\lambda = \phi(x_{2i}'\beta_2)/\Phi(x_{2i}'\beta_2)$ is Heckman's lambda. If $\sigma_{12} \neq 0$ then OLS estimation of (5.4) produces biased results. Therefore, the null hypothesis of $\sigma_{12} = 0$ should be tested for checking the adequacy of using the sample selection model.

Omitted variables are very important issue for a sample selection model. As it is not clear how x_{1i} and x_{2i} differ "some sensitivity analysis with respect to the imposed exclusion restrictions should be performed to make sure that the λ term is not incorrectly picking up the effect of omitted variables" (Verbeek, 2000, p. 211).

Maximum likelihood method is used for estimation of parameter vectors β_1 and β_2 in (5.4) and (5.5). The Huber-White-sandwich estimator of variance is used in place of the traditional calculation to avoid the problem of heteroscedasticity (for details on obtaining robust variance estimates see Stata User Manual, [U]23.11).

In the sample selection model described above, the dependent variable is a continuous one. As the data on household arrears are only reported to lie in certain intervals, the continuous variable was obtained in the following way: the middle of the interval to which the size of household arrears belongs was divided by the number of members in the household. Although measurement error present in such a constructed dependent variable does not affect the validity of estimation and inference procedures it requires additional caution in interpreting estimation results presented in section 5.2.4.

5.2. Estimation and results discussion

5.2.1. Presence of household arrears: logit model

Table 5.1 shows the estimation results of logit model, in which dependent variable is equal to one if a household has arrears for housing and utilities, otherwise it is zero. Almost all coefficients are highly significant and have reasonable economic interpretation.

Table 5.1. Logit model estimation results for the presence of household arrears

Dependent variable: presence of household arrears				
Independent variables:	Coefficient	p-value	Marginal effect	p-value for ME
Income per member	-0.3578	0.000	-0.067	0.000
Income per member squared	0.0166	0.031	0.003	0.031
Share of wage arrears in income	0.6878	0.008	0.128	0.008
Presence of wage arrears	0.3743	0.000	0.073	0.000
Workers of a state-owned firm	-0.0159	0.789	-0.003	0.788
Worklessness status (pensioner household is base category)				
all-work	0.4619	0.000	0.087	0.000
mix-work	0.5788	0.000	0.117	0.000
workless	0.7210	0.000	0.152	0.000
Age of household head	-0.0097	0.000	-0.002	0.000
Sex of household head	-0.1262	0.016	-0.023	0.016
Education (base category is high school and no education)				
higher, basic higher, unfinished higher	-0.0733	0.330	-0.013	0.324
vocational technical, specialized secondary	0.0678	0.269	0.013	0.272
Location (town is base category)				
city	0.4634	0.000	0.089	0.000
rural	-0.3328	0.000	-0.060	0.000
Flat is privatized	-0.3146	0.000	-0.061	0.000
Flat characteristics				
living area	0.0001	0.771	0.000	0.771
heating	0.3155	0.000	0.059	0.000
gas	1.1429	0.000	0.199	0.000
water	0.4533	0.000	0.082	0.000
hot water	-0.0399	0.597	-0.007	0.595
Constant	-1.2976	0.000		
Pseudo R2 = 0.1874, Prob > chi2 = 0.0000				

Complete estimation output from Stata for all models presented in this chapter is given in Appendix C.

The coefficients for both household income and its square are significant. Initially, this estimated quadratic function of income decreases reaching minimum at UAH 10.8 thousand and then increases. As the income per member greater than the minimizing value of income rarely occurs in the sample (16 observations from 9318), it is possible to conclude that the probability of household arrears presence falls in income per member and an increase in income of a rich household affects the probability of arrears less than the same absolute increase in income of a poor household.

The coefficients of the share of wage arrears in yearly household income and the dummy for the presence of wage arrears are positive and significant. The effect of the size of wage arrears on the probability of having arrears for housing and utilities seems to be of smaller importance than the effect of the presence of wage arrears. The presence of any small amount of wage arrears increases the probability of having arrears by more than 0.07 on average (see marginal effects in table 5.1), while the increase in the share of wage arrears in income necessary to produce the same effect on the probability must be more than 0.54.

The fact that some household members work at an organization financed from the state budget has no significant effect on the probability of having arrears for housing and utilities although budget enterprises were encouraged to perform mutual cancellation of wage and household arrears since 1999.

The more unsuccessful the experience of a household on the labor market the higher the probability of being in arrears. The probability of arrears for mix-work households (households in which some working age members are out of work) is 34% higher than for all-work households (households with no workless members). For workless households (households in which all

working age members are out of work) the probability of arrears is 74% larger than for all-work households.

An interesting fact is that the older the household head the smaller the probability of being in arrears. This is also confirmed by the result that pensioner households have smaller probability of being in arrears (pensioner household is the base category for worklessness status dummy variables). One possible cause of this result is that a lot of old people (veterans and participants of World War II, the families of those killed at the time of World War II, etc.) got additional privileges on payments for housing and utilities. Another hypothesis is that older people have higher moral costs of being in debt.

The negative significant coefficient for a household head being male could reflect that women-headed households earn less income. This model also shows that the education level of the household head has statistically insignificant effect on the probability of the presence of household arrears.

Living in a city raises the probability of having arrears for communal and housing services compared to living in a town. The explanation could be that it is easier to avoid informal social control and enforcement in large densely populated areas. Living in the countryside decreases the probability of arrears perhaps because the value of services per capita consumed in rural areas is much less than in a town or a city.

The fact that a household flat is privatized significantly reduces the probability of arrears. However, it cannot be argued from this that there is such a casual effect of the possibility of enforcement on the probability of arrears other things being equal because flat privatization can be endogenous to household arrears. The coefficient for the dummy for the privatized flat can certainly show only that households with privatized flats have lower probabilities of being in arrears.

The last subset of independent variables reported in table 5.1 includes factors that determine the size of payment for the housing and communal services. These factors are the area of the flat and the presence of heating, gas, and cold and hot water. It can be seen from the table that major determinants of the services value, such as presence of gas or heating, have positive significant effect on the probability of having arrears.

5.2.2. Persistence of household arrears: logit model

Table 5.2 presents the estimation results of logit model, in which the dependent variable is equal to one if a household has arrears of more than UAH 1000 (the highest range of household arrears reported in the data set available); otherwise it is zero.

Table 5.2. Logit model estimation results for the persistence of household arrears

Dependent variable: household arrears are more than 1000 hr (last range reported)				
Independent variables:	Coefficient	p-value	Marginal effect	p-value for ME
Income per member	-0.470	0.000	-0.0060	0.000
Income per member squared	0.018	0.000	0.0002	0.000
Share of wage arrears in income	0.450	0.156	0.0057	0.162
Presence of wage arrears	0.291	0.037	0.0040	0.057
Workers of a state-owned organization	-0.300	0.008	-0.0037	0.009
Worklessness status (pensioner household is base category)				
all work	0.768	0.000	0.0106	0.001
mixed	1.257	0.000	0.0249	0.000
all out of work	1.411	0.000	0.0336	0.000
Age of household head	-0.015	0.000	-0.0002	0.001
Sex of household head	0.061	0.565	0.0008	0.567
Education (base category is high school and no education)				
higher, basic higher, unfinished higher	-0.285	0.045	-0.0034	0.033
vocational technical training, specialized secondary	-0.293	0.017	-0.0036	0.014
Location (town is base category)				
city	0.623	0.000	0.0088	0.000
rural	-1.140	0.000	-0.0130	0.000
Flat is privatized	-0.643	0.000	-0.0100	0.000
Flat characteristics				
living area	0.002	0.000	0.0000	0.001
heating	1.459	0.000	0.0227	0.000
gas	0.721	0.001	0.0088	0.002
water	0.142	0.628	0.0018	0.617
hot water	0.169	0.203	0.0022	0.226
Constant	-3.827	0.000		
Pseudo R2 = 0.2417, Prob > chi2 = 0.0000				

The households that have arrears of more than 1000 UAH can be considered as persistent nonpayers as on average they have debt for more than 1 year. Theoretically, a household could be paying off the arrears when it is observed and still has a large stock of arrears. However, such situation seems to be unlikely, as the incentive structure for compliance has not recently been changed in Ukraine. Moreover, the microlevel data collected by PADCO for some rayons (PADCO, 1998, p.10) show that the customers can mainly be divided into two categories: those who pay with some delay and persistent nonpayers. Thus, this logit model is assumed to analyze the probability of being a persistent non-payer.

As tables 5.1 and 5.2 show the determinants of having arrears and being a persistent non-payer are almost the same. However, there are some important differences.

First, vocational technical training or specialized secondary education and higher education have almost the same statistically significant negative effect on the persistence of household arrears. This means that people with low level of education (high school or no education) and, perhaps, little human capital, are more likely to be persistent non-payers.

The second difference is that the coefficient for the share of wage arrears is not significant. This tends to confirm the hypothesis that the fact of the presence of wage arrears is more important determinant of household arrears than the size of wage arrears.

Finally, the fact that a household member works at a state-owned organization has significant negative effect on the probability of being a persistent nonpayer. If an individual works at an organization which is financed from the State Budget he or she is able to mutually cancel wage and household arrears. Thus, a positive effect of the government program for the mutual cancellation of wage and household arrears on compliance shows up in the data.

5.2.3. Size of household arrears: ordered probit

Table 5.3 presents the estimation results of ordered probit model for the size of household arrears. The comparison of the table 5.1, 5.2, and 5.3 shows that the determinants of the presence, the persistence, and the size of the household arrears are almost the same. There are two differences worth mentioning.

Table 5.3. Ordered probit model estimation results for the size of household arrears

Dependent variable: range of the size of household arrears per household member		
Independent variables:	Coefficient	p-value
Income per member	-0.147	0.000
Income per member squared	0.007	0.005
Share of wage arrears in income	0.572	0.000
Presence of wage arrears	0.130	0.001
Workers of a state-owned firm	-0.056	0.072
Worklessness status (pensioner household is base category)		
all work	0.166	0.001
mixed	0.214	0.000
all out of work	0.415	0.000
Age of household head	-0.007	0.000
Sex of household head	-0.091	0.001
Education (base category is high school and no education)		
higher, basic higher, unfinished higher	-0.072	0.079
vocational technical training, specialized secondary	-0.016	0.632
Location (town is base category)		
city	0.308	0.000
rural	-0.195	0.000
Flat is privatized	-0.245	0.000
Flat characteristics		
living area	0.000	0.399
heating	0.322	0.000
gas	0.630	0.000
water	0.230	0.000
hot water	0.007	0.873
Pseudo R2 = 0.0901, Prob > chi2 = 0.0000		

First, the size of wage arrears is more important than their presence in explaining the size of household arrears, while in explaining the presence and persistence of household arrears the opposite observation was made. Another difference is that only higher education has significant negative effect on the size of household arrears.

5.2.4. Size of household arrears: tobit II

The estimation results of Heckman's sample selection model (tobit II model) are presented in tables 5.4, 5.5, and 5.6.

Table 5.4. Heckman's sample selection model estimation results for the size of household arrears per household member (the selection equation)

Dependent variable: presence of household arrears		
Independent variables:	Coefficient	p-value
Income per member	-0.20	0.000
Income per member squared	0.01	0.001
Share of wage arrears in income	0.41	0.007
Presence of wage arrears	0.22	0.000
Workers of a state-owned organization	-0.01	0.863
Worklessness status (pensioner household is base category)		
all-work	0.25	0.000
mix-work	0.32	0.000
workless	0.41	0.000
Age of household head	-0.01	0.000
Sex of household head	-0.08	0.010
Education (base category is high school and no education)		
higher, basic higher, unfinished higher	-0.04	0.335
vocational technical training, specialized secondary	0.04	0.318
Location (town is base category)		
city	0.28	0.000
rural	-0.18	0.000
Flat is privatized	-0.20	0.000
Flat characteristics		
living area	0.00	0.752
heating	0.19	0.000
gas	0.67	0.000
water	0.26	0.000
hot water	-0.02	0.588
Constant	-0.76	0.000

It can be noted from the table 5.4 that the equation for selection of being in arrears is similar to the logit model reported in section 5.2.1, as it should be.

However, the equation for the size of arrears per household member presented in table 5.5 gives us interesting results that are completely different from those obtained from the models discussed earlier in this chapter.

Table 5.5. Heckman's sample selection model estimation results for the size of household arrears per household member (the arrears size equation)

Dependent variable: proxy for the size of household arrears per household member		
Independent variables:	Coefficient	p-value
Income per member	29.54	0.004
Income per member squared	-1.66	0.055
Share of wage arrears in income	214.83	0.000
Presence of wage arrears	-36.27	0.005
Workers of a state-owned firm	-43.87	0.000
Worklessness status (pensioner household is base category)		
all-work	-74.50	0.001
mix-work	-81.17	0.000
workless	35.91	0.196
Age of household head	-1.46	0.003
Sex of household head	-15.36	0.171
Education (base category is high school and no education)		
higher, basic higher, unfinished higher	-32.24	0.057
vocational technical training, specialized secondary	-42.14	0.002
Location (town is base category)		
city	57.83	0.000
rural	-22.98	0.049
Flat is privatized	-65.51	0.000
Flat characteristics		
living area	0.21	0.334
heating	123.05	0.000
gas	31.38	0.025
water	-7.26	0.583
hot water	17.57	0.288
Constant	286.36	0.000

Although the selection equation still shows that households with lower income are more likely to have arrears than those with higher income, the effect of income on the size of household arrears per member is positive and

significant for those households that have nonzero arrears for housing and utilities.

This seems to be one of those cases in which using a sample selection model "has led to reinterpretation of earlier results" (Greene, 2000, p. 927).

There may be several explanations for this. A likely explanation could be the following. For the households wealthy enough, the payments for housing and utilities are imperceptible and they are least likely to be in arrears. For other households, the lower is the income of a household, the higher is the share of services value paid for the household by the local government in the form of the housing subsidy. Thus, poor households are well motivated to pay on time and to gradually pay off old debts as it makes them eligible for the subsidy. This tends to confirm the fact reported in (PADCO, 2001) that the targeted subsidies program improved compliance. At the same time, a bit richer households have lower incentives to comply as they are eligible for no or very small subsidy. In order to create right incentives for these richer households and to change the positive relationship between the size of arrears and income detected on both micro and macro levels, fines for late payments should be reintroduced.

Heckman's sample selection model also sheds some more light on the role of the household worklessness status in explaining the size of household arrears. For those households that are in arrears, the fact that a household is of the all-work type has almost the same effect as the fact that a household is of the mix-work type (the Wald test does not reject the hypothesis of the coefficients equality). Workless households have higher arrears than the households of other types do. However, this effect of worklessness is not statistically different from the effect of the fact that a household is of pensioner type (a pensioner household is the base category for dummy variables for household worklessness status).

Estimation results of Heckman's sample selection model provide no other qualitative changes to the analysis of the determinants of the size of household arrears based on the ordered probit model estimation discussed in the previous section.

Table 5.6. Heckman's sample selection model estimation results for the size of household arrears per household member (additional output and test statistics).

	Estimates:	p-value
Ath(rho)	-0.09443	0.000
Ln(sigma)	5.665915	0.000
rho	-0.09415	
sigma	288.8521	
lambda	-27.1947	
Wald test of indep. eqns. (rho = 0): chi2(1) = 12.96 Prob > chi2 = 0.0003		

As can be seen from the table 5.6, the Wald test rejects the hypothesis of the independence of the selection equation and the arrears size equation (p-value is 0.0003). On one hand, modeling results should be taken with caution since only proxies for the size of arrears are used in the estimation. On the other hand, the estimation results are not sensitive to a particular version of the proxies (see section 4.1 for details on constructing proxies for the size of arrears). Moreover, the main estimation results are not sensitive to a particular model specification. A number of different specifications were examined. This inspires confidence in the correctness of the results obtained.

The estimation results for all microeconometric models discussed in this chapter are given in Appendix D in the form convinient for the comparison.

Chapter 6

CONCLUSIONS AND POLICY IMPLICATIONS

Under the central planning, the housing and utilities sector was mainly financed by the government. In 1995, Ukraine abandoned subsidizing housing and utilities for the whole population. Instead, targeted housing subsidies for the needy were introduced. The speed and the magnitude of the adjustment in the housing and utilities sector were striking. Household arrears for housing and utilities became a self-protection mechanism against the huge increase in prices for housing and utilities accompanied by the population impoverishment and widespread wage arrears.

Even after the economic situation in Ukraine has stabilized and the restructuring of the housing and utilities sector has mainly occurred, the government continues to grant the "implicit subsidy" to households in the form of the contract unenforceability in the field and the absence of fines for late payments. The reason for that seems to be the lack of political will and the consequences of that can be seen even at the macro level. Visual and regression analysis of the regional level data shows that while there is an anticipated positive relationship between household and wage arrears, income is positively related to the size of household arrears for housing and utilities, which at first may seem as an unexpected result. Microeconometric analysis corroborates these findings and reveals some other determinants of household arrears.

Estimation results at the household level show that the poor are more likely to be in arrears. However, for those households that have nonzero arrears for housing and utilities, the size of arrears is positively related to income. The most likely explanation for this could be the following. The households wealthy enough have no problem with payments for housing and utilities.

Poor households are well motivated to pay on time and to gradually pay off old debts by the government program of housing subsidies. Slightly richer households have lower incentives to comply as they are eligible for either no or very small subsidy. In order to create right incentives for these slightly richer households and to change the positive relationship between the size of arrears and income detected on both micro and macro levels, fines for late payments should be reintroduced.

The size and the probability of arrears are negatively related to the age of the household head. This is also confirmed by the result that pensioner households have smaller probability of having arrears. One possible cause of this result is that a lot of old people got additional privileges on payments for housing and utilities. Another psychological explanation stating that older people have higher moral costs of being in arrears points towards the expediency of reintroducing fines.

The micro and macro level studies confirmed that arrears to household play a very important role explaining arrears for utilities and housing. Therefore, collecting arrears and fines should be conditional on the presence and the size of wage, pension, and social payment arrears. A positive effect of the government program for the mutual cancellation of wage and household arrears on compliance shows up in the data.

Living in a city raises the probability of having arrears for communal and housing services compared to living in a town. The explanation could be that it is easier to avoid informal social control and enforcement in large densely populated areas. Living in the countryside decreases the probability of arrears perhaps because the value of services per capita consumed in rural areas is less than in a town or a city.

The fact that a household flat is privatized significantly reduces the probability and the size of arrears, perhaps, because payments for services can be collected through the court from the owners of privatized flats. Besides,

households which possess privatized flats are more likely to be wealthy and to have no problems with payments for housing and utilities.

If the head of a household is male the probability and the size of arrears are lower. This could reflect the fact that women-headed households earn less income. Households headed by people with low level of education and, perhaps, little human capital are more likely to be persistent non-payers.

The more unsuccessful the experience of a household as a whole on the labor market, the higher the probability and the size of arrears. Thus, collecting fines and arrears should be means tested, so as not to punish those workless and poor households a second time.

Weak contract enforcement and nonpayments have taken deep roots in the Ukrainian economy. It would be unfair and not so efficient to fight against household arrears alone. The decision to reintroduce fines for late payments and collecting arrears from households should be adopted simultaneously with making providers of housing and utilities responsible for the regularity and the quality of the services. Wage arrears are another closely related issue. Wages must be one of the most important employer's liabilities as it is in developed and some transition countries. Creating economic laws punishing the breach of contract in all fields of society's economic life is necessary for successful continuation of reforms in the country and should be a priority task for Ukrainian legislators.

WORKS CITED

Alfandari, G. and Schaffer, M. E. 1996. 'Arrears' in the Russian Enterprise Sector. In *Enterprise Restructuring and Economic Policy in Russia*, ed. Simon Commander, Qimiao Fan, and Mark E. Schaffer, Washington, DC: EDI/World Bank.

Bagratian, H. and Gurgen, E. 1997. Payments arrears in the gas and electric power sectors of the Russian Federation and Ukraine. IMF Working Paper No. WP/97/162.

Coudouel, A., McAuley, A., and Micklewright, J. 1997. Transfers and Exchange between households in Uzbekistan. In *Household welfare in Central Asia*, ed. Falkingham, J., Klugman, S., and Micklewright, J. London, Macmillan.

Daianu, D., 1997. Structure and strain in explaining inter-enterprise arrears. The Davidson Institute Working Paper Series. Working paper number 97.

Earle, J. S. and Sabirianova, K. Z., 2000. Equilibrium Wage Arrears: A Theoretical and Empirical Analysis of Institutional Lock-In. Institute for the Study of Labor (IZA), Series IZA Discussion Papers, No. 196.

Greene, W. 2000. Econometric analysis. New Jersey, Prentice-Hall, Inc.

Gregg, P., and Wadsworth, J. 2001. Everything you ever wanted to know about measuring worklessness and polarization at the household level but were afraid to ask. Forthcoming Oxford Bulletin of Economics and Statistics.

Gryshyna, O. 2000. The incidence of wage arrears in Ukraine: evidence from enterprise and household level data. Kyiv-Mohyla Academy University, EERC MA Program in Economics. MA Thesis.

Lehmann, H., Wadsworth, J., and Acquisiti, A.,. 1999. Grime and Punishment: Job Insecurity and Wage Arrears in the Russian Federation. *Journal of Comparative Economics 27*: 595-617.

Morozov, A. Pinto, B. Drebentsov, V. 2000. Dismantling Russia's Nonpayments System: Creating Conditions for Growth. World Bank Technical Paper No. 471.

PADCO. 1998. Arrears for housing and communal services in 1996 – 1998. PADCO / USAID Analytic Report Series on Issues of Target Social Aid in Ukraine. Report No. 26.

PADCO. 2001. Arrears for housing and communal services in 2000. PADCO / USAID Analytic Report Series on Issues of Tariff Reform and Communal Enterprises Restructuring. Report No.1.

PADCO. 2001b. The history of tariff reform in Ukraine. PADCO / USAID Analytic Report Series on Issues of Tariff Reform and Communal Enterprises Restructuring. Report No.2.

State Committee of Statistics of Ukraine. 1998-2001. On payment for housing and communal services. Monthly express-reports.

State Committee of Statistics of Ukraine. 1998-2001.b. Wages and state of their payment. Monthly express-reports.

State Committee of Statistics of Ukraine. 2000. Statistical Yearbook of Ukraine.

Verbeek, M.. 2000. A guide to modern econometrics. Chichester,: John Wiley and Sons, Ltd.

UNDP, 1996. Ukraine Human Development Report. UNDP, Kyiv, June 1996. Chpater 7. Shelter.

Appendix A

SHORT DESCRIPTION OF THE HOUSEHOLD SURVEY DATASETS

A1. Brief description of SPSS microfile for individual level data

- Type of locality (city, town, or rural area).
- Oblast.
- Sex.
- Family status.
- Family relation with head of household.
- Level of education.
- Area of education.
- Working experience in years.
- Socio-economic status (employee, employer, self-employed, housekeeper, student, pensioner, child).
- Ownership type of enterprise of principal place of business (state enterprise or organization, private, joint stock company, …).
- Industry of principal place of business (from 26 industries).
- Type of independent labor activity (trade, construction works, … - 25 types).
- Pension type (disability, social, old-age, …).
- Wage at principal place of business.
- Additional payments at principal place of business.
- Wage in-kind.
- Wage at nonprincipal place of business.
- Income from entrepreneurial activity.
- Income from independent labor activity.
- Pension.

- Stipend.
- Unemployment benefits.
- Other income.
- Weight, height.
- Health selfevaluation, presence of different kinds of diseases.
- Data on smoking.
- Number of children born.
- Sum of wage, pension, stipend, and social payment arrears to a household member.

A.2. Brief description of SPSS microfiles for household level data

- Number of the household members.
- Number of children in the household by age intervals.
- Number of the old-aged and the disabled in the household.
- Age, sex, education level, and socio-economic status of the household head.
- Expenditures on food bought by 52 types of food.
- Value of food produced in the household 52 types of food.
- Value of food received as a gift by 52 types of food.
- Expenditure on transportation, personal services, education, taxes, … (51 types).
- Value of subsidies and privileges from government.
- Household income by 18 types of sources.
- Data on housing: type, ownership type, area, rooms, age, availability of different facilities.
- Data on land which is in use by the household: area, use type,…
- Data on the household livestock.
- Sum of household arrears for housing and communal services, gas, electricity, and etc.

- Selfevaluation of household income (enough even for savings, enough but not for savings, not enough,…).
- Frequency of availability of different kinds of food.
- Availability of different kinds of consumer goods.
- Amount of money which interviewee needs not to feel himself or herself poor.
- Change in financial position of a household during last year (has improved, worsened,…).
- Expectations about household financial position in the next year.
- Household expectations about economic perspectives of Ukraine in the next year.

Appendix B

LISTINGS OF COMPUTER PROGRAMS

B1. SQL computer program for creating household worklessness status variable

```
SELECT CODE_FAM,
CASE
 WHEN SUM(CASE WHEN SES_MEM = 9 OR SES_MEM = 10 THEN 1 ELSE 0 END) = 0
  THEN CASE WHEN
   SUM(CASE WHEN SES_MEM <= 5 OR SES_MEM = 7 THEN 1 ELSE 0 END) != 0
   THEN 3 ELSE 1 END
  ELSE CASE WHEN
   SUM(CASE WHEN SES_MEM <= 5 OR SES_MEM = 7 THEN 1 ELSE 0 END) = 0
   THEN 2  ELSE 4 END
END
AS POLAR_STATUS
INTO POLAR_RES
FROM MEMB_DB
GROUP BY CODE_FAM
ORDER BY CODE_FAM ASC
```

B2. Stata computer program for preparing data, building cross tables, and microeconometric modelling

```
log using "C:\EERC\Thesis\stata\log1.smcl", replace
insheet using C:\EERC\Thesis\stata\hs_reduced_2.dat
set more off
gen totres1000=totalres/1000
gen sqtotres1000 = totres1000*totres1000
gen sm_slry1000 = sm_slry/1000
gen sh_warrs = sm_slry / totalres
gen membinc1000= membinc/1000
gen sqmeminc1000= membinc1000*membinc1000
gen edu123 = 0
replace  edu123=1 if ( edu_1==1 | edu_2==1 | edu_3==1 )
gen edu45 = 0
replace  edu45=1 if ( edu_4==1 | edu_5==1 )
replace heating = heating / hsize
replace gascentr = gascentr / hsize
replace runwater = runwater / hsize
replace  hotwater =  hotwater / hsize
gen harrs_pc=0
```

```
replace harrs_pc = 50/hsize if sm_srv==1

replace harrs_pc = 200/hsize if sm_srv==2

replace harrs_pc = 400/hsize if sm_srv==3

replace harrs_pc = 750/hsize if sm_srv==4

replace harrs_pc = 2000/hsize if sm_srv==5

gen srv_val =  exp27_1 + exp27_2 + exp27_3

regress srv_val sliv heating gascentr runwater hotwater city  rural
membinc if  db_srv ~= 1

predict srv_val1

gen harrs_term = 0

replace harrs_term = 50/srv_val1 if sm_srv==1

replace harrs_term = 200/srv_val1 if sm_srv==2

replace harrs_term = 400/srv_val1 if sm_srv==3

replace harrs_term = 750/srv_val1 if sm_srv==4

replace harrs_term = 2000/srv_val1 if sm_srv==5

gen subs_elig = 0

replace subs_elig = 1 if (srv_val1/cashinc) > 0.25

gen db12_srv = 0

replace db12_srv = 1 if sm_srv==5


set matsize 200


logit db_srv  membinc1000 sqmeminc1000 sh_warrs db_slry stateown all_empl
all_mix all_unem age_head sex_head edu123 edu45 city rural prvtzd sliv
heating gascentr runwater hotwater, robust

mfx compute


logit db12_srv  membinc1000 sqmeminc1000 sh_warrs db_slry stateown all_empl
all_mix all_unem age_head sex_head edu123 edu45 city rural prvtzd sliv
heating gascentr runwater hotwater, robust

mfx compute


oprobit harrs_pc  membinc1000 sqmeminc1000 sh_warrs db_slry stateown
all_empl all_mix all_unem age_head sex_head edu123 edu45 city rural prvtzd
sliv heating gascentr runwater hotwater, robust


heckman harrs_pc  membinc1000 sqmeminc1000 sh_warrs db_slry stateown
all_empl all_mix all_unem age_head sex_head edu123 edu45 city rural prvtzd
sliv heating gascentr runwater hotwater, select(db_srv = membinc1000
sqmeminc1000 sh_warrs db_slry stateown all_empl all_mix all_unem age_head
sex_head edu123 edu45 city rural prvtzd sliv heating gascentr runwater
hotwater) robust


test  all_empl = all_mix
```

```
/* Adding interactions of arrears to household and worklessness*/


gen aw_arr=0
replace aw_arr = 1 if ( all_empl==1 & db_slry==1 )
gen mw_arr=0
replace mw_arr = 1 if ( all_mix==1 & db_slry==1 )
gen nw_arr=0
replace nw_arr = 1 if ( all_unem==1 & db_slry==1 )
gen p_arr=0
replace p_arr = 1 if ( all_unem==0 & all_empl==0 & all_mix==0 & db_slry==1)


logit db_srv membinc1000 sqmeminc1000 sh_warrs stateown p_arr all_empl
aw_arr all_mix mw_arr all_unem nw_arr age_head sex_head edu123 edu45 city
rural prvtzd sliv heating gascentr runwater hotwater, robust
mfx compute


/* Cross tables*/


tabulate  db_srv [fweight = w_q], summarize(membinc)
tabulate  harrs_pc [fweight = w_q], summarize(membinc)
tabulate  db_srv [fweight = w_q], summarize(db_slry)
tabulate  db_srv [fweight = w_q], summarize(sh_warrs)


tabulate  db_srv [fweight = w_q], summarize(all_empl)
tabulate  db_srv [fweight = w_q], summarize(all_mix)
tabulate  db_srv [fweight = w_q], summarize(all_unem)
gen pen_hsld = 0
replace pen_hsld = 1 if (all_unem==0 & all_empl==0 & all_mix==0)
tabulate  db_srv [fweight = w_q], summarize(pen_hsld)
```

Appendix C

STATA LOG FILE

```
log:  C:\EERC\Thesis\stata\log1.smcl
  log type:  smcl
 opened on:  19 May 2002, 13:43:54

. insheet using C:\EERC\Thesis\stata\hs_reduced_2.dat
(96 vars, 9318 obs)

. set more off

. gen totres1000=totalres/1000

. gen sqtotres1000 = totres1000*totres1000

. gen sm_slry1000 = sm_slry/1000

. gen sh_warrs = sm_slry / totalres

. gen membinc1000= membinc/1000

. gen sqmeminc1000= membinc1000*membinc1000

. gen edu123 = 0

. replace  edu123=1 if ( edu_1==1 | edu_2==1 | edu_3==1 )
(1749 real changes made)

. gen edu45 = 0

. replace  edu45=1 if ( edu_4==1 | edu_5==1 )
(2955 real changes made)

. replace heating = heating / hsize
(3165 real changes made)

. replace gascentr = gascentr / hsize
(4509 real changes made)

. replace runwater = runwater / hsize
(4729 real changes made)

. replace  hotwater =  hotwater / hsize
(2277 real changes made)

. gen harrs_pc=0

. replace harrs_pc = 50/hsize if sm_srv==1
(643 real changes made)

. replace harrs_pc = 200/hsize if sm_srv==2
(805 real changes made)

. replace harrs_pc = 400/hsize if sm_srv==3
(434 real changes made)

. replace harrs_pc = 750/hsize if sm_srv==4
(494 real changes made)

. replace harrs_pc = 2000/hsize if sm_srv==5
(464 real changes made)

. gen srv_val =  exp27_1 + exp27_2 + exp27_3

. regress srv_val sliv heating gascentr runwater hotwater city  rural  membinc if
db_srv ~= 1
```

```
      Source |       SS       df       MS              Number of obs =    6478
-------------+------------------------------           F(  8,  6469) =  840.58
       Model |   597185954     8  74648244.2           Prob > F      =  0.0000
    Residual |   574486970  6469  88806.1478           R-squared     =  0.5097
-------------+------------------------------           Adj R-squared =  0.5091
       Total |  1.1717e+09  6477  180897.472           Root MSE      =  298.00

------------------------------------------------------------------------------
     srv_val |      Coef.   Std. Err.      t    P>|t|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
        sliv |   .2388576   .0678292     3.52   0.000     .1058898    .3718253
     heating |    60.1236   14.29975     4.20   0.000     32.09136    88.15583
    gascentr |   400.3674   8.893883    45.02   0.000     382.9325    417.8024
    runwater |   103.1051    10.1679    10.14   0.000      83.1727    123.0376
    hotwater |   153.2335   14.24815    10.75   0.000     125.3024    181.1646
        city |   19.77426   11.04383     1.79   0.073    -1.875298    41.42382
       rural |  -39.80818   10.00772    -3.98   0.000    -59.42662   -20.18974
     membinc |   .0292972   .0029807     9.83   0.000      .023454    .0351405
       _cons |   58.79848    11.0965     5.30   0.000     37.04567     80.5513
------------------------------------------------------------------------------

. predict srv_val1
(option xb assumed; fitted values)

. gen harrs_term = 0

. replace harrs_term = 50/srv_val1 if sm_srv==1
(643 real changes made)

. replace harrs_term = 200/srv_val1 if sm_srv==2
(805 real changes made)

. replace harrs_term = 400/srv_val1 if sm_srv==3
(434 real changes made)

. replace harrs_term = 750/srv_val1 if sm_srv==4
(494 real changes made)

. replace harrs_term = 2000/srv_val1 if sm_srv==5
(464 real changes made)

. gen subs_elig = 0

. replace subs_elig = 1 if (srv_val1/cashinc) > 0.25
(2690 real changes made)

. gen db12_srv = 0

. replace db12_srv = 1 if sm_srv==5
(464 real changes made)

.
. set matsize 200

.
. logit db_srv  membinc1000 sqmeminc1000 sh_warrs db_slry stateown all_empl all_mix
all_unem age_head sex_head edu123 edu45 city rural prvtzd sliv heat
> ing gascentr runwater hotwater, robust

Iteration 0:   log likelihood = -5729.3163
Iteration 1:   log likelihood = -4728.1116
Iteration 2:   log likelihood = -4657.8147
Iteration 3:   log likelihood = -4655.5802
Iteration 4:   log likelihood = -4655.5766

Logit estimates                                   Number of obs   =       9318
                                                  Wald chi2(20)   =    1644.96
                                                  Prob > chi2     =     0.0000
Log likelihood = -4655.5766                       Pseudo R2       =     0.1874

------------------------------------------------------------------------------
             |               Robust
      db_srv |      Coef.   Std. Err.      z    P>|z|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
 membinc1000 |  -.3577886   .0637539    -5.61   0.000    -.4827439   -.2328334
sqmeminc1000 |   .0166179   .0076946     2.16   0.031     .0015368    .0316991
    sh_warrs |   .6877752   .2593868     2.65   0.008     .1793864    1.196164
     db_slry |   .3742971   .0788474     4.75   0.000     .2197591     .528835
```

```
    stateown |  -.0159499   .0595234    -0.27   0.789    -.1326136    .1007138
    all_empl |     .46189   .0896286     5.15   0.000     .2862212    .6375589
     all_mix |   .5788016    .099629     5.81   0.000     .3835324    .7740708
    all_unem |   .7209883   .1089919     6.62   0.000     .5073682    .9346085
    age_head |  -.0096858   .0021405    -4.53   0.000    -.0138811   -.0054905
    sex_head |  -.1261771   .0523134    -2.41   0.016    -.2287095   -.0236447
      edu123 |  -.0732726   .0751906    -0.97   0.330    -.2206435    .0740982
       edu45 |   .0677563   .0612667     1.11   0.269    -.0523242    .1878368
        city |   .4633555   .0646957     7.16   0.000     .3365543    .5901567
       rural |  -.3327834   .0832232    -4.00   0.000    -.4958979   -.1696689
      prvtzd |  -.3145873   .0632738    -4.97   0.000    -.4386018   -.1905729
        sliv |   .0001369    .000471     0.29   0.771    -.0007861     .00106
     heating |   .3155481   .0865835     3.64   0.000     .1458474    .4852487
    gascentr |   1.142928   .0715013    15.98   0.000     1.002788    1.283068
    runwater |    .453324   .0825659     5.49   0.000     .2914978    .6151503
    hotwater |  -.0399135   .0754044    -0.53   0.597    -.1877034    .1078764
       _cons |  -1.297555   .1943182    -6.68   0.000    -1.678412   -.9166983
------------------------------------------------------------------------------

. mfx compute

Marginal effects after logit
      y  = Pr(db_srv) (predict)
         =  .24698072
------------------------------------------------------------------------------
variable |      dy/dx    Std. Err.     z    P>|z|  [    95% C.I.   ]      X
---------+--------------------------------------------------------------------
mem~1000 |   -.066542      .01183   -5.62   0.000  -.089729 -.043355   1.99455
sqm~1000 |   .0030906      .00143    2.16   0.031   .000285  .005896   5.53008
sh_warrs |   .1279133      .04826    2.65   0.008   .033328  .222498   .043679
 db_slry*|   .0728795      .01596    4.57   0.000   .041595  .104164   .243400
stateown*|  -.0029633      .01105   -0.27   0.788  -.024613  .018686   .369822
all_empl*|   .0873812       .0171    5.11   0.000   .053866  .120897   .420691
 all_mix*|   .1168556      .02141    5.46   0.000   .074895  .158816   .188131
all_unem*|   .1520029      .02512    6.05   0.000   .102777  .201229   .094441
age_head |  -.0018014       .0004   -4.52   0.000  -.002582 -.001021   54.0279
sex_head*|   -.023419       .0097   -2.41   0.016  -.042426 -.004412   .469414
  edu123*|  -.0134688      .01366   -0.99   0.324  -.040233  .013295   .187701
   edu45*|     .01268      .01153    1.10   0.272  -.009926  .035286   .317128
    city*|   .0889378      .01281    6.94   0.000   .063826  .114049   .355978
   rural*|   -.060313      .01462   -4.12   0.000  -.088977 -.031649   .351685
  prvtzd*|  -.0611825      .01286   -4.76   0.000  -.086396 -.035969   .795450
    sliv |   .0000255      .00009    0.29   0.771  -.000146  .000197   41.2651
 heating*|   .0594177       .0165    3.60   0.000    .02707  .091765   .418223
gascentr*|   .1994887      .01149   17.36   0.000   .176971  .222006   .600021
runwater*|   .0818097      .01439    5.68   0.000   .053601  .110018   .625456
hotwater*|  -.0073926      .01391   -0.53   0.595  -.034653  .019868   .296737
------------------------------------------------------------------------------
(*) dy/dx is for discrete change of dummy variable from 0 to 1

.
. logit db12_srv  membinc1000 sqmeminc1000 sh_warrs db_slry stateown all_empl all_mix
all_unem age_head sex_head edu123 edu45 city rural prvtzd sliv he
> ating gascentr runwater hotwater, robust

Iteration 0:   log likelihood = -1844.1665
Iteration 1:   log likelihood = -1616.6861
Iteration 2:   log likelihood = -1421.0187
Iteration 3:   log likelihood = -1400.7568
Iteration 4:   log likelihood = -1398.5944
Iteration 5:   log likelihood = -1398.5207
Iteration 6:   log likelihood = -1398.5205

Logit estimates                                   Number of obs   =       9318
                                                  Wald chi2(20)   =     633.85
                                                  Prob > chi2     =     0.0000
Log likelihood = -1398.5205                       Pseudo R2       =     0.2417

------------------------------------------------------------------------------
             |               Robust
    db12_srv |      Coef.   Std. Err.      z    P>|z|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
 membinc1000 |  -.4699196   .0872611    -5.39   0.000    -.6409483    -.298891
sqmeminc1000 |   .0175633    .003648     4.81   0.000     .0104134    .0247132
    sh_warrs |    .449656   .3172375     1.42   0.156    -.1721181     1.07143
     db_slry |    .291363   .1393645     2.09   0.037     .0182135    .5645124
    stateown |  -.3000487   .1138601    -2.64   0.008    -.5232105    -.076887
    all_empl |   .7676948   .2161565     3.55   0.000     .3440359    1.191354
```

```
     all_mix |     1.2572   .2305382     5.45   0.000     .8053532    1.709046
    all_unem |   1.411301   .2321805     6.08   0.000     .9562359    1.866367
    age_head |  -.0146973   .0039684    -3.70   0.000    -.0224753   -.0069194
    sex_head |   .0613308   .1066869     0.57   0.565    -.1477716    .2704332
      edu123 |  -.2853509   .1424586    -2.00   0.045    -.5645646   -.0061373
       edu45 |  -.2932164   .1223016    -2.40   0.017    -.5329232   -.0535096
        city |   .6232949   .1317625     4.73   0.000     .3650452    .8815446
       rural |  -1.139714   .3155126    -3.61   0.000    -1.758107   -.5213204
      prvtzd |  -.6429822   .1095287    -5.87   0.000    -.8576545   -.4283099
        sliv |   .0019513    .000554     3.52   0.000     .0008655     .003037
     heating |   1.459331    .250225     5.83   0.000     .9688995    1.949763
    gascentr |      .7211    .211232     3.41   0.001     .3070929    1.135107
    runwater |   .1424466   .2939367     0.48   0.628    -.4336588     .718552
    hotwater |   .1686236   .1323538     1.27   0.203    -.0907851    .4280324
       _cons |  -3.826837   .4210021    -9.09   0.000    -4.651986   -3.001688
------------------------------------------------------------------------------

. mfx compute

Marginal effects after logit
      y  = Pr(db12_srv) (predict)
         =  .01294256
------------------------------------------------------------------------------
variable |      dy/dx    Std. Err.     z    P>|z|  [    95% C.I.   ]      X
---------+--------------------------------------------------------------------
mem~1000 |  -.0060032       .0012   -4.99   0.000  -.008363 -.003643   1.99455
sqm~1000 |   .0002244      .00005    4.48   0.000   .000126  .000323   5.53008
sh_warrs |   .0057444      .00411    1.40   0.162  -.002312  .013801   .043679
 db_slry*|   .0040161      .00211    1.90   0.057  -.000126  .008158   .243400
stateown*|  -.0037028      .00143   -2.60   0.009  -.006497 -.000909   .369822
all_empl*|    .010642      .00325    3.27   0.001   .004268  .017016   .420691
 all_mix*|   .0248657      .00677    3.67   0.000   .011599  .038133   .188131
all_unem*|   .0336051      .00931    3.61   0.000   .015349  .051861   .094441
age_head |  -.0001878      .00006   -3.36   0.001  -.000297 -.000078   54.0279
sex_head*|   .0007851      .00137    0.57   0.567  -.001903  .003474   .469414
  edu123*|  -.0033525      .00157   -2.13   0.033  -.006431 -.000274   .187701
   edu45*|  -.0035669      .00145   -2.46   0.014  -.006411 -.000723   .317128
    city*|   .0088186      .00234    3.77   0.000   .004228  .013409   .355978
   rural*|  -.0129774      .00275   -4.72   0.000  -.018365 -.007589   .351685
  prvtzd*|   -.010034       .0024   -4.17   0.000  -.014746 -.005322   .795450
    sliv |   .0000249      .00001    3.40   0.001   .000011  .000039   41.2651
 heating*|   .0226644      .00545    4.16   0.000    .01199  .033338   .418223
gascentr*|   .0087601      .00277    3.16   0.002   .003324  .014197   .600021
runwater*|   .0017897      .00358    0.50   0.617  -.005222  .008801   .625456
hotwater*|   .0022297      .00184    1.21   0.226  -.001377  .005837   .296737
------------------------------------------------------------------------------
(*) dy/dx is for discrete change of dummy variable from 0 to 1

.
. oprobit harrs_pc  membinc1000 sqmeminc1000 sh_warrs db_slry stateown all_empl
all_mix all_unem age_head sex_head edu123 edu45 city rural prvtzd sliv
> heating gascentr runwater hotwater, robust

Iteration 0:   log likelihood = -14154.796
Iteration 1:   log likelihood = -12918.269
Iteration 2:   log likelihood = -12879.157
Iteration 3:   log likelihood = -12878.842
Iteration 4:   log likelihood = -12878.841

Ordered probit estimates                          Number of obs   =       9318
                                                  Wald chi2(20)   =    2371.60
                                                  Prob > chi2     =     0.0000
Log likelihood = -12878.841                       Pseudo R2       =     0.0901

------------------------------------------------------------------------------
             |               Robust
    harrs_pc |      Coef.   Std. Err.      z    P>|z|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
 membinc1000 |  -.1467102   .0248656    -5.90   0.000     -.195446   -.0979744
sqmeminc1000 |   .0067672   .0024333     2.78   0.005      .001998    .0115363
    sh_warrs |   .5716476   .1168857     4.89   0.000     .3425559    .8007394
     db_slry |   .1296773   .0383482     3.38   0.001     .0545161    .2048385
    stateown |  -.0560852   .0312244    -1.80   0.072     -.117284    .0051135
    all_empl |   .1664681   .0483785     3.44   0.001      .071648    .2612881
     all_mix |   .2143213    .052997     4.04   0.000      .110449    .3181936
    all_unem |   .4145637    .058614     7.07   0.000     .2996823    .5294451
    age_head |  -.0068374   .0011411    -5.99   0.000    -.0090739   -.0046008
    sex_head |  -.0906503   .0277538    -3.27   0.001    -.1450468   -.0362538
```

```
      edu123 |  -.0715974   .0408245    -1.75   0.079     -.151612    .0084171
       edu45 |  -.0156813   .0327137    -0.48   0.632     -.079799    .0484364
        city |   .3084541   .0345537     8.93   0.000     .2407302    .3761781
       rural |  -.1946805   .0412651    -4.72   0.000    -.2755587   -.1138023
      prvtzd |  -.2449717     .03415    -7.17   0.000    -.3119046   -.1780389
        sliv |   .0002376   .0002818     0.84   0.399    -.0003147      .00079
     heating |   .3217007   .0461492     6.97   0.000       .23125    .4121515
    gascentr |   .6295336   .0373548    16.85   0.000     .5563195    .7027478
    runwater |   .2301354   .0411108     5.60   0.000     .1495598     .310711
    hotwater |   .0065311   .0410078     0.16   0.873    -.0738426    .0869049
-------------+----------------------------------------------------------------
       _cut1 |   .6611488   .1023405          (Ancillary parameters)
       _cut2 |   .6630401   .1023594
       _cut3 |   .6641759   .1023695
       _cut4 |   .6706256   .1024022
       _cut5 |   .6912935   .1024468
       _cut6 |     .73219   .1024394
       _cut7 |   .7869465   .1024773
       _cut8 |   .7873525   .1024813
       _cut9 |   .8551414   .1024571
      _cut10 |   .8568473   .1024821
      _cut11 |   .8662676   .1025647
      _cut12 |   .8848486   .1026621
      _cut13 |   1.029943   .1023553
      _cut14 |   1.031394   .1023555
      _cut15 |   1.144164   .1026232
      _cut16 |     1.1618   .1026442
      _cut17 |   1.337703   .1029255
      _cut18 |   1.338897   .1029286
      _cut19 |   1.344876   .1029768
      _cut20 |   1.416426   .1031376
      _cut21 |   1.430235   .1032322
      _cut22 |   1.430866   .1032264
      _cut23 |   1.507658   .1034597
      _cut24 |   1.686407   .1033698
      _cut25 |   1.688018   .1033883
      _cut26 |   1.816721   .1036431
      _cut27 |   1.821248   .1036659
      _cut28 |   1.833095    .103745
      _cut29 |   1.960085   .1041599
      _cut30 |   2.071137    .104255
      _cut31 |   2.256849   .1052381
      _cut32 |   2.528808   .1066789
      _cut33 |   2.705309   .1082217
      _cut34 |   3.243196   .1190076
------------------------------------------------------------------------------

.
. heckman harrs_pc  membinc1000 sqmeminc1000 sh_warrs db_slry stateown all_empl
all_mix all_unem age_head sex_head edu123 edu45 city rural prvtzd sliv
> heating gascentr runwater hotwater, select(db_srv = membinc1000 sqmeminc1000
sh_warrs db_slry stateown all_empl all_mix all_unem age_head sex_head ed
> u123 edu45 city rural prvtzd sliv heating gascentr runwater hotwater) robust

Iteration 0:   log likelihood = -24764.747
Iteration 1:   log likelihood = -24763.717
Iteration 2:   log likelihood = -24763.645
Iteration 3:   log likelihood = -24763.645

Heckman selection model                         Number of obs      =      9318
(regression model with sample selection)        Censored obs       =      6478
                                                Uncensored obs     =      2840

                                                Wald chi2(20)      =    523.96
Log likelihood = -24763.64                      Prob > chi2        =    0.0000

------------------------------------------------------------------------------
             |               Robust
             |      Coef.   Std. Err.      z    P>|z|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
harrs_pc     |
 membinc1000 |   29.54475   10.31672     2.86   0.004     9.324346    49.76515
sqmeminc1000 |  -1.660703   .8640067    -1.92   0.055    -3.354125     .032719
    sh_warrs |   214.8326   46.81788     4.59   0.000     123.0713     306.594
     db_slry |  -36.27099    12.9944    -2.79   0.005    -61.73955   -10.80244
    stateown |  -43.86669   11.27373    -3.89   0.000     -65.9628   -21.77059
    all_empl |  -74.50105   22.04439    -3.38   0.001    -117.7073   -31.29484
     all_mix |  -81.17265   22.63441    -3.59   0.000    -125.5353   -36.81003
```

```
    all_unem |   35.91376   27.76098     1.29   0.196    -18.49675    90.32428
    age_head |  -1.459287    .488044    -2.99   0.003    -2.415835   -.5027379
    sex_head |  -15.36112   11.22546    -1.37   0.171    -37.36263    6.640383
      edu123 |  -32.23981   16.91647    -1.91   0.057    -65.39548    .9158658
       edu45 |  -42.14104   13.64564    -3.09   0.002    -68.88601   -15.39608
        city |   57.82601   12.48799     4.63   0.000     33.35001    82.30202
       rural |  -22.98224   11.68839    -1.97   0.049    -45.89107   -.0734044
      prvtzd |  -65.50832   13.24132    -4.95   0.000    -91.46084    -39.5558
        sliv |   .2063973   .2138316     0.97   0.334     -.212705    .6254996
     heating |   123.0541    17.2863     7.12   0.000      89.1736    156.9347
    gascentr |   31.38159   13.98313     2.24   0.025     3.975155    58.78802
    runwater |  -7.257631   13.22748    -0.55   0.583    -33.18302    18.66776
    hotwater |   17.57032   16.54271     1.06   0.288    -14.85279    49.99344
       _cons |   286.3585   44.03601     6.50   0.000     200.0495    372.6675
-------------+----------------------------------------------------------------
db_srv       |
 membinc1000 |  -.1989992   .0270152    -7.37   0.000     -.251948   -.1460504
sqmeminc1000 |   .0089681   .0026704     3.36   0.001     .0037342     .014202
    sh_warrs |   .4058849   .1511694     2.68   0.007     .1095984    .7021714
     db_slry |   .2150909   .0459927     4.68   0.000     .1249468     .305235
    stateown |    -.00607   .0352495    -0.17   0.863    -.0751577    .0630177
    all_empl |   .2492215   .0520479     4.79   0.000     .1472095    .3512335
     all_mix |   .3182724   .0579527     5.49   0.000     .2046873    .4318576
    all_unem |   .4067868   .0629194     6.47   0.000     .2834671    .5301066
    age_head |  -.0057427   .0012628    -4.55   0.000    -.0082177   -.0032677
    sex_head |  -.0794076   .0307489    -2.58   0.010    -.1396743   -.0191408
      edu123 |  -.0427739   .0443689    -0.96   0.335    -.1297353    .0441874
       edu45 |    .036196   .0362382     1.00   0.318    -.0348296    .1072217
        city |   .2834754   .0385745     7.35   0.000     .2078709      .35908
       rural |  -.1830649   .0466453    -3.92   0.000     -.274488   -.0916419
      prvtzd |  -.1970834   .0378682    -5.20   0.000    -.2713037   -.1228632
        sliv |   .0000917   .0002904     0.32   0.752    -.0004775    .0006609
     heating |   .1936478   .0518108     3.74   0.000     .0921005    .2951951
    gascentr |   .6676185   .0394139    16.94   0.000     .5903687    .7448683
    runwater |   .2635524   .0463935     5.68   0.000     .1726229     .354482
    hotwater |  -.0248006    .045818    -0.54   0.588    -.1146021     .065001
       _cons |  -.7561421   .1114167    -6.79   0.000    -.9745149   -.5377694
-------------+----------------------------------------------------------------
     /athrho |  -.0944272   .0262306    -3.60   0.000    -.1458383    -.043016
    /lnsigma |   5.665915   .0340876   166.22   0.000     5.599104    5.732725
-------------+----------------------------------------------------------------
         rho |  -.0941475   .0259981                     -.1448131   -.0429895
       sigma |   288.8521   9.846277                      270.1843    308.8097
      lambda |   -27.1947   7.636214                     -42.16141     -12.228
------------------------------------------------------------------------------
Wald test of indep. eqns. (rho = 0): chi2(1) =    12.96   Prob > chi2 = 0.0003
------------------------------------------------------------------------------


. 
. test  all_empl = all_mix

 ( 1)  [harrs_pc]all_empl - [harrs_pc]all_mix = 0.0

           chi2(  1) =    0.36
         Prob > chi2 =    0.5499


. 
. 
. /* Adding interactions of arrears to household and worklessness*/
. 
. gen aw_arr=0

. replace aw_arr = 1 if ( all_empl==1 & db_slry==1 )
(1421 real changes made)

. gen mw_arr=0

. replace mw_arr = 1 if ( all_mix==1 & db_slry==1 )
(652 real changes made)

. gen nw_arr=0

. replace nw_arr = 1 if ( all_unem==1 & db_slry==1 )
(110 real changes made)

. gen p_arr=0
```

```
. replace p_arr = 1 if ( all_unem==0 & all_empl==0 & all_mix==0 & db_slry==1)
(85 real changes made)

.
. logit db_srv membinc1000 sqmeminc1000 sh_warrs stateown p_arr all_empl aw_arr
all_mix mw_arr all_unem nw_arr age_head sex_head edu123 edu45 city rura
> l prvtzd sliv heating gascentr runwater hotwater, robust

Iteration 0:   log likelihood = -5729.3163
Iteration 1:   log likelihood = -4726.6437
Iteration 2:   log likelihood = -4656.0918
Iteration 3:   log likelihood = -4653.8349
Iteration 4:   log likelihood = -4653.8312

Logit estimates                                   Number of obs   =       9318
                                                  Wald chi2(23)   =    1648.98
                                                  Prob > chi2     =     0.0000
Log likelihood = -4653.8312                       Pseudo R2       =     0.1877

------------------------------------------------------------------------------
             |               Robust
      db_srv |      Coef.   Std. Err.      z    P>|z|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
 membinc1000 |  -.3578652   .0642684    -5.57   0.000     -.483829   -.2319014
sqmeminc1000 |   .0167092   .0077651     2.15   0.031     .0014899    .0319284
    sh_warrs |   .6593366   .2570497     2.57   0.010     .1555285    1.163145
    stateown |  -.0143318   .0596203    -0.24   0.810    -.1311854    .1025218
       p_arr |   .6061263   .3139041     1.93   0.053    -.0091145    1.221367
    all_empl |   .4608313   .0932938     4.94   0.000     .2779787    .6436838
      aw_arr |   .4111572   .0918298     4.48   0.000     .2311741    .5911402
     all_mix |   .5957245   .1083543     5.50   0.000      .383354     .808095
      mw_arr |   .3621368   .1227985     2.95   0.003     .1214561    .6028175
    all_unem |    .781925   .1142678     6.84   0.000     .5579641    1.005886
      nw_arr |  -.0146457   .2336383    -0.06   0.950    -.4725685     .443277
    age_head |   -.009584   .0021458    -4.47   0.000    -.0137897   -.0053783
    sex_head |  -.1263654   .0523431    -2.41   0.016    -.2289559   -.0237749
      edu123 |  -.0736248   .0752038    -0.98   0.328    -.2210215     .073772
       edu45 |   .0680949   .0613368     1.11   0.267     -.052123    .1883128
        city |   .4632384   .0647598     7.15   0.000     .3363116    .5901652
       rural |    -.33028   .0832295    -3.97   0.000    -.4934068   -.1671532
      prvtzd |  -.3128692   .0633233    -4.94   0.000    -.4369807   -.1887578
        sliv |   .0001156   .0004678     0.25   0.805    -.0008012    .0010324
     heating |   .3184473   .0867122     3.67   0.000     .1484946    .4884001
    gascentr |   1.145586   .0715429    16.01   0.000     1.005364    1.285807
    runwater |    .451669   .0825818     5.47   0.000     .2898116    .6135264
    hotwater |  -.0397715   .0755224    -0.53   0.598    -.1877927    .1082498
       _cons |  -1.317214   .1953429    -6.74   0.000    -1.700079    -.934349
------------------------------------------------------------------------------

. mfx compute

Marginal effects after logit
      y  = Pr(db_srv) (predict)
         =  .24687488
------------------------------------------------------------------------------
variable |      dy/dx    Std. Err.      z    P>|z|  [    95% C.I.   ]       X
---------+--------------------------------------------------------------------
mem~1000 |   -.066537      .01192   -5.58   0.000  -.089903 -.043172   1.99455
sqm~1000 |   .0031067      .00144    2.15   0.031   .000277  .005937   5.53008
sh_warrs |   .1225889      .04781    2.56   0.010   .028887  .216291   .043679
stateown*|  -.0026622      .01106   -0.24   0.810  -.024345  .019021   .369822
   p_arr*|   .1282318      .07334    1.75   0.080  -.015507  .271971   .009122
all_empl*|   .0871537       .0178    4.90   0.000   .052269  .122038   .420691
  aw_arr*|   .0817486      .01933    4.23   0.000    .04387  .119627   .152501
 all_mix*|   .1204927      .02338    5.15   0.000   .074664  .166322   .188131
  mw_arr*|   .0724381      .02616    2.77   0.006   .021171  .123705   .069972
all_unem*|   .1661639      .02666    6.23   0.000   .113902  .218426   .094441
  nw_arr*|  -.0027132      .04313   -0.06   0.950  -.087238  .081812   .011805
age_head |  -.0017819       .0004   -4.47   0.000  -.002564    -.001   54.0279
sex_head*|  -.0234471       .0097   -2.42   0.016  -.042459 -.004435   .469414
  edu123*|  -.0135288      .01365   -0.99   0.322  -.040287  .013229   .187701
   edu45*|   .0127401      .01154    1.10   0.270  -.009887  .035367   .317128
    city*|   .0888906      .01282    6.93   0.000   .063764  .114017   .355978
   rural*|  -.0598535      .01462   -4.09   0.000  -.088514 -.031193   .351685
  prvtzd*|  -.0608181      .01287   -4.73   0.000  -.086034 -.035603   .795450
    sliv |   .0000215      .00009    0.25   0.805  -.000149  .000192   41.2651
 heating*|   .0599532      .01653    3.63   0.000   .027556   .09235   .418223
gascentr*|   .1998628      .01151   17.37   0.000   .177306  .222419   .600021
```

```
runwater*|    .081496      .01439    5.66   0.000   .053285  .109707   .625456
hotwater*|  -.0073643      .01393   -0.53   0.597   -.03466  .019932   .296737
------------------------------------------------------------------------------
(*) dy/dx is for discrete change of dummy variable from 0 to 1

.
. /* Cross tables*/
.
. tabulate  db_srv [fweight = w_q], summarize(membinc)

            |               Summary of MEMBINC
     DB_SRV |        Mean   Std. Dev.       Freq.        Obs.
------------+------------------------------------------------
          0 |   2067.5769   1270.0996    12002831    12002831
          1 |   1790.1507   1163.1373     5676727     5676727
------------+------------------------------------------------
      Total |   1978.4982   1243.5282    17679558    17679558

. tabulate  harrs_pc [fweight = w_q], summarize(membinc)

            |               Summary of MEMBINC
   harrs_pc |        Mean   Std. Dev.       Freq.        Obs.
------------+------------------------------------------------
          0 |   2067.5769   1270.0996    12002831    12002831
       6.25 |   1453.1853   517.73775        6769        6769
   7.142857 |   1176.7911   646.81018        3631        3631
   8.333333 |   1406.8584   530.70106       36160       36160
         10 |   1463.3068   703.74368      114953      114953
       12.5 |   1437.3016   924.70106      205459      205459
   16.66667 |    1773.544   1036.5167      247978      247978
   18.18182 |    1457.907           0         974         974
         25 |   1900.4928   1011.9736      285302      285302
   28.57143 |   1083.0703    249.4377        8026        8026
   33.33333 |   1305.7074   473.61942       36249       36249
         40 |   1477.8268   807.97397       86638       86638
         50 |   1869.5339   1144.8493      634573      634573
   57.14286 |   1621.6733   444.32026        6785        6785
   66.66666 |   1919.9558   1164.9137      434984      434984
         80 |   1573.9276   647.47995       72588       72588
        100 |   1908.5411   1258.8743      620712      620712
   107.1429 |   1188.8411   404.78982        3368        3368
        125 |   1349.8274   611.01507       24753       24753
   133.3333 |   1687.7471   743.99409      242838      242838
        150 |    1277.857    428.4036       49115       49115
   166.6667 |   402.89084           0        1435        1435
      187.5 |   1611.6941   1409.0771      278064      278064
        200 |   2049.7716   1186.3318      457613      457613
   222.2222 |    429.7443   144.89844        5028        5028
        250 |   1832.3454   1291.0754      314005      314005
   285.7143 |   1233.1123   662.45707       12983       12983
   333.3333 |   1029.3825   531.96122       31219       31219
        375 |   1918.5756   1457.6203      259792      259792
        400 |   2023.6891   1753.5362      199241      199241
        500 |   1393.6761   706.70176      306621      306621
   666.6667 |    1684.259   933.19088      302625      302625
        750 |   2212.9562    1240.384      118667      118667
       1000 |   1853.7109   1300.3494      195444      195444
       2000 |   2062.7317   1219.9899       72135       72135
------------+------------------------------------------------
      Total |   1978.4982   1243.5282    17679558    17679558

. tabulate  db_srv [fweight = w_q], summarize(db_slry)

            |               Summary of DB_SLRY
     DB_SRV |        Mean   Std. Dev.       Freq.        Obs.
------------+------------------------------------------------
          0 |   .21706121   .41224465    12002831    12002831
          1 |    .3025208   .45934954     5676727     5676727
------------+------------------------------------------------
      Total |   .24450142   .42979121    17679558    17679558

. tabulate  db_srv [fweight = w_q], summarize(sh_warrs)

            |               Summary of sh_warrs
     DB_SRV |        Mean   Std. Dev.       Freq.        Obs.
------------+------------------------------------------------
          0 |   .03553354   .10905604    12002831    12002831
          1 |   .06188198   .15927611     5676727     5676727
```

```
------------+--------------------------------------------------
      Total |   .04399376   .12795099    17679558    17679558

.
. tabulate  db_srv [fweight = w_q], summarize(all_empl)

           |               Summary of ALL_EMPL
    DB_SRV |        Mean   Std. Dev.       Freq.        Obs.
-----------+--------------------------------------------------
         0 |   .40009294   .48991693    12002831    12002831
         1 |   .50644764   .49995847     5676727     5676727
-----------+--------------------------------------------------
      Total |   .43424236   .49565709    17679558    17679558

. tabulate  db_srv [fweight = w_q], summarize(all_mix)

           |               Summary of ALL_MIX
    DB_SRV |        Mean   Std. Dev.       Freq.        Obs.
-----------+--------------------------------------------------
         0 |   .17587309   .38071217    12002831    12002831
         1 |   .23765931   .42564938     5676727     5676727
-----------+--------------------------------------------------
      Total |   .19571202   .39674782    17679558    17679558

. tabulate  db_srv [fweight = w_q], summarize(all_unem)

           |               Summary of ALL_UNEM
    DB_SRV |        Mean   Std. Dev.       Freq.        Obs.
-----------+--------------------------------------------------
         0 |   .08309115   .27601995    12002831    12002831
         1 |   .10928533    .3119969     5676727     5676727
-----------+--------------------------------------------------
      Total |   .09150184   .28832144    17679558    17679558

. gen pen_hsld = 0

. replace pen_hsld = 1 if (all_unem==0 & all_empl==0 & all_mix==0)
(2765 real changes made)

. tabulate  db_srv [fweight = w_q], summarize(pen_hsld)

           |               Summary of pen_hsld
    DB_SRV |        Mean   Std. Dev.       Freq.        Obs.
-----------+--------------------------------------------------
         0 |   .34094282   .47402619    12002831    12002831
         1 |   .14660772   .35371446     5676727     5676727
-----------+--------------------------------------------------
      Total |   .27854378   .44828245    17679558    17679558

.
.
end of do-file

. exit, clear
```

Appendix D

SUMMARY OF ESTIMATION RESULTS FOR ALL MICROECONOMETRIC MODELS

Table D.1. Summary of estimation results for all microeconometric models

	Presence: logit (ME)	Persistence: logit (ME)	Size: oprobit	Size: tobit II
Income per member	-0.067	-0.0060	-0.147	29.54
Income per member squared	0.003	0.0002	0.007	-1.66(-)
Share of wage arrears in income	0.128	0.0057(--)	0.572	214.83
Presence of wage arrears	0.073	0.0040(-)	0.130	-36.27
Workers of a state-owned firm	-0.003(--)	-0.0037	-0.056(-)	-43.87
Worklessness status (pensioner household is base category)				
all-work	0.087	0.0106	0.166	-74.50
mix-work	0.117	0.0249	0.214	-81.17
workless	0.152	0.0336	0.415	35.91(--)
Age of household head	-0.002	-0.0002	-0.007	-1.46
Sex of household head	-0.023	0.0008(--)	-0.091	-15.36(--)
Education (base category is high school and no education)				
higher education	-0.013(--)	-0.0034	-0.072(-)	-32.24(-)
vocational technical, specialized secondary	0.013(--)	-0.0036	-0.016(--)	-42.14
Location (town is base category)				
city	0.089	0.0088	0.308	57.83
rural	-0.060	-0.0130	-0.195	-22.98
Flat is privatized	-0.061	-0.0100	-0.245	-65.51
Flat characteristics				
living area	0.000(--)	0.0000	0.000(--)	0.21(--)
heating	0.059	0.0227	0.322	123.05
gas	0.199	0.0088	0.630	31.38
water	0.082	0.0018(--)	0.230	-7.26(--)
hot water	-0.007(--)	0.0022(--)	0.007(--)	17.57(--)
Constant				286.36

no marks - significant at 5%, (-) significant at 10%, (--) not significant at 10%